

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 10:50:47 ; Search time 71.1206 Seconds
(without alignments)
1073.493 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSVKDSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	2512	99.4	481	19	AAW59045		Human MNTF1-F3 pro
2	2485	98.3	729	22	AAU20459		Human secreted pro
3	2485	98.3	729	22	AAU20616		Human secreted pro
4	1848	73.1	975	22	ABB61030		Drosophila melanog
5	870	34.4	858	22	AAM79209		Human protein SEQ
6	870	34.4	858	22	AAM80193		Human protein SEQ
7	863	34.2	858	8	AAP70099		Sequence of elonga
8	863	34.2	858	8	AAP70100		Sequence of varian
9	844.5	33.4	699	22	ABB67414		Drosophila melanog
10	844.5	33.4	832	22	ABB58973		Drosophila melanog
11	807.5	32.0	842	21	AAV43637		Amino acid sequenc
12	807.5	32.0	842	21	AAV43638		Amino acid sequenc
13	806	31.9	884	24	ABJ25942		Aspergillus fumiga
14	743	29.4	845	23	AAB71315		L. major 4G2-83 ex
15	743	29.4	845	23	AAU71861		Leishmania antigen
16	726.5	28.7	750	21	AAG51496		Arabidopsis thalia
17	726.5	28.7	784	21	AAG51495		Arabidopsis thalia
18	726.5	28.7	805	21	AAG51494		Arabidopsis thalia
19	554.5	21.9	361	19	AAW70256		Leishmania antigen
20	554.5	21.9	361	23	AAB71306		L. major CD4+T cel
21	554.5	21.9	361	23	AAE24970		Leishmania major 4
22	554.5	21.9	361	23	ABG60908		Leishmania antigen
23	554.5	21.9	361	23	AAU71852		Leishmania antigen
24	460	18.2	860	22	ABB60832		Drosophila melanog
25	448	17.7	1120	23	AAU99609		Human glioma antig
26	448	17.7	1122	22	AAU20462		Human secreted pro
27	406.5	16.1	1044	23	ABP73793		Candida albicans e
28	395.5	15.7	857	22	AAB94484		Human protein sequ
29	395.5	15.7	857	24	ABJ19771		Human MP21 protein
30	387.5	15.3	1013	24	ABJ25752		Aspergillus fumiga
31	385.5	15.3	735	22	AAB96105		Putative P. abyssi
32	379	15.0	1087	24	ABJ26352		Aspergillus fumiga
33	325	12.9	543	21	AAB07853		Amino acid sequenc
34	291	11.5	630	22	AAU20467		Human secreted pro
35	284.5	11.3	268	21	AAB43153		Human ORFX ORF2917
36	232.5	9.2	150	23	ABP07738		Human ORFX protein
37	219	8.7	695	23	ABB48403		Listeria monocytog
38	212	8.4	268	23	ABB57238		Mouse ischaemic co
39	206.5	8.2	707	23	ABP65900		Bifidobacterium lo
40	200	7.9	336	21	AAV95044		Candida albicans p
41	188	7.4	693	22	AAU33742		Staphylococcus aur
42	188	7.4	693	24	ABJ18971		Pathogen specific
43	188	7.4	715	22	AAU36850		Staphylococcus aur
44	188	7.4	715	22	AAU37143		Staphylococcus aur
45	187.5	7.4	693	19	AAW80722		S. pneumoniae elon

ALIGNMENTS

RESULT 1

AAW59045

ID AAW59045 standard; Protein; 481 AA.

XX

AC AAW59045;

XX

DT 11-AUG-1998 (first entry)

XX

DE Human MNTF1-F3 protein fragment.

XX

KW Motoneuronotrophic factor; MNTF-1; MNTF1-F3; human; axon regeneration;
KW motoneurone; diagnose; treatment; disease; wound healing; scar tissue;
KW keloid.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 194

FT /note= "encoded by TTC"

FT Misc-difference 248

FT /note= "encoded by AGC"

FT Misc-difference 461

FT /note= "encoded by ACY"

XX

XX

PN WO9813492-A2.

XX

PD 02-APR-1998.

XX

PF 22-SEP-1997; 97WO-US17142.

XX

PR 12-SEP-1997; 97US-0928862.

PR 27-SEP-1996; 96US-0026792.

PR 15-NOV-1996; 96US-0751225.

XX

PA (KMBI-) KM BIOTECH INC.

XX

PI Chau RMW;

XX

DR WPI; 1998-230703/20.

DR N-PSDB; AAV11747.

XX

PT Motoneurotrophic factor MNTF1-F3 and MNTF1-F6 - useful for
PT motoneuron regeneration, diagnosing or treating motoneuron disease
PT and to accelerate wound healing without scar formation

XX

PS Claim 3; Fig 2A; 78pp; English.

XX

CC This sequence represents a novel human motoneurotrophic factor, MNTF1-F3.
CC Such factors are used to promote regeneration of the axon of a
CC motoneurone, to diagnose and treat motoneurone disease in a mammal or to
CC accelerate wound healing whilst concomitantly minimising or inhibiting
CC scar tissue and/or keloid formation in an area associated with a wound.
CC For promoting axonal regeneration, the polypeptide is administered at a
CC concentration of 5 ng-50 mg, whereas for inhibiting hereditary
CC motoneurone disease, the dosage is 5-100 (especially 30-50) ng per kg
CC body weight.

XX

SQ Sequence 481 AA;

Query Match 99.4%; Score 2512; DB 19; Length 481;
Best Local Similarity 99.4%; Pred. No. 1.8e-235;
Matches 478; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 RAYYSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWGDIY 60
         |||
Db      1 RAYYSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWGDIY 60

Qy     61 FNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKREL 120
         |||
Db     61 FNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKREL 120

Qy    121 KLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPSQKVGAKPKIEHTYTGGVSDSLGEAMSDC 180
         |||
Db    121 KLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPSQKVGAKPKIEHTYTGGVSDSLGEAMSDC 180

Qy    181 DPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVG 240
         |||
Db    181 DPDGPLMCHTTKMYSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSPICPVG 240

Qy    241 RLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVI 300
         |||
Db    241 RLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVI 300

Qy    301 KIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKY 360
         |||
Db    301 KIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKY 360

Qy    361 SEIDIKVADPVVTFCEVTVETSSSLKCFEATPNKKNKITMIAEPLKGLAEDIENEVVQIT 420
         |||
Db    361 SEIDIKVADPVVTFCEVTVETSSSLKCFEATPNKKNKITMIAEPLKGLAEDIENEVVQIT 420

Qy    421 WNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQ 480
         |||
Db    421 WNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQ 480

Qy    481 G 481
         |
Db    481 G 481
```

RESULT 2

AAU20459

ID AAU20459 standard; Protein; 729 AA.

XX

AC AAU20459;

XX

DT 06-DEC-2001 (first entry)

XX

DE Human secreted protein, Seq ID No 451.

XX

KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;

KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;

KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
 KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
 KW Gaucher's disease; neurological disease; cerebrovascular disorder;
 KW thrombosis; wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO200155326-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01347.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451931/48.
 DR N-PSDB; AAS33168.
 XX
 PT New nucleic acids and polypeptides, useful for diagnosing, preventing
 PT or treating medical conditions -
 XX
 PS Claim 11; SEQ ID No 451; 753pp; English.
 XX
 CC The invention relates to novel isolated nucleic acid molecules (I)
 CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
 CC the prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression of secreted proteins. (I) and complementary
 CC sequences may also be used as DNA probes in diagnostic assays (e.g.
 CC polymerase chain reactions (PCR)) to detect and quantitate the presence
 CC of similar nucleic acid sequences in samples, and so which patients may
 CC be in need of restorative therapy. (II) may also be used as antigens in
 CC the production of antibodies and in assays to identify modulators
 CC (agonists and antagonists) of the expression and activity of the secreted
 CC proteins. The anti-(II) antibodies and antagonists may also be used to
 CC down regulate expression and activity of (II). The anti-(II) antibodies
 CC may also be used as diagnostic agents for detecting the presence of (II)
 CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
 CC disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
 CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
 CC angina and thrombosis), infections caused by bacteria, viruses and
 CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),
 CC agonists, antagonists and antibodies can also be used to promote wound
 CC healing, maintain organs before transplantation, and support cell culture
 CC of primary tissues. AAU20342-AAU20666 represent human secreted protein
 CC amino acid sequences, and related sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed

CC specification but was obtained in electronic format directly from WIPO
CC at: ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 729 AA;

Query Match 98.3%; Score 2485; DB 22; Length 729;
Best Local Similarity 99.2%; Pred. No. 1.5e-232;
Matches 474; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY      4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDIYFNP 63
      |||
Db      49 YSTDENLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDIYFNP 108

QY     64 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKEELKLN 123
      |||
Db     109 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTXEELKLN 168

QY     124 IRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 183
      |||
Db     169 IRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 228

QY     184 GPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 243
      |||
Db     229 GPLMCHTTKMYSTDDGVQFHA FGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 288

QY     244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
      |||
Db     289 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 348

QY     304 VEPVNPSELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 363
      |||
Db     349 VEPVNPSELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 408

QY     364 DIKVADPVVTFCE TVVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNR 423
      |||
Db     409 DIKVADPVVTFCE TVVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNR 468

QY     424 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
      |||
Db     469 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 526
```

RESULT 3

AAU20616

ID AAU20616 standard; Protein; 729 AA.

XX

AC AAU20616;

XX

DT 06-DEC-2001 (first entry)

XX

DE Human secreted protein, Seq ID No 608.

XX

KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;

KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;

KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;

KW multiple sclerosis; cancer; hyperproliferative disorder; infection;

KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.

XX
OS Homo sapiens.

XX
PN WO200155326-A2.

XX
PD 02-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01347.

XX
PR 31-JAN-2000; 2000US-0179065.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Rosen CA, Barash SC, Ruben SM;

XX
DR WPI; 2001-451931/48.

XX
DR N-PSDB; AAS33325.

XX
PT New nucleic acids and polypeptides, useful for diagnosing, preventing
PT or treating medical conditions -

XX
PS Claim 11; SEQ ID No 608; 753pp; English.

XX
CC The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
CC angina and thrombosis), infections caused by bacteria, viruses and
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),
CC agonists, antagonists and antibodies can also be used to promote wound
CC healing, maintain organs before transplantation, and support cell culture
CC of primary tissues. AAU20342-AAU20666 represent human secreted protein
CC amino acid sequences, and related sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification but was obtained in electronic format directly from WIPO
CC at: [ftp.wipo.int/pub/published_pct_sequences](ftp:wipo.int/pub/published_pct_sequences).

XX

SQ Sequence 729 AA;

Query Match 98.3%; Score 2485; DB 22; Length 729;
Best Local Similarity 99.2%; Pred. No. 1.5e-232;
Matches 474; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy      4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWGDYFNP 63
         |||
Db      49 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWGDYFNP 108

Qy      64 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKHEELKLN 123
         |||
Db     109 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTXEELKLN 168

Qy     124 IRPLRLRVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 183
         |||
Db     169 IRPLRLRVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 228

Qy     184 GPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEEDSQICTVGRLW 243
         |||
Db     229 GPLMCHTTKMYSTDDGVQFHAFAFGRVLSGTIHAGQPVKVLGENYTLEDEEDSQICTVGRLW 288

Qy     244 ISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
         |||
Db     289 ISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 348

Qy     304 VEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI 363
         |||
Db     349 VEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI 408

Qy     364 DIKVADPVVTFCEVTVETSSSLKCFAPNKKNKITMIAEPLKGLAEDIENEVVQITWNR 423
         |||
Db     409 DIKVADPVVTFCEVTVETSSSLKCFAPNKKNKITMIAEPLKGLAEDIENEVVQITWNR 468

Qy     424 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
         |||
Db     469 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 526
```

RESULT 4

ABB61030

ID ABB61030 standard; Protein; 975 AA.

XX

AC ABB61030;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 9882.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

```

PD 27-SEP-2001.
XX
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL05133.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 9882; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 975 AA;

```

QY 303 AVEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSE 362
 ||||| :|||
 Db 594 AVEPVNPSELPMKLDGLRKVNKSYPLLSTRVEESGEHVILGTGELYLDCVMHDLRKMYSE 653
 QY 363 IDIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAEPLEKGLAEDIENEV VQITWN 422
 ||||| :|||
 Db 654 IDIKVADPVVAFCE TVVETSSSLKCF AETPNKKNKITMISEPLEKGLAEDIENGTV CINWN 713
 QY 423 RKKLG EFFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKALLG SVKDSIVQG 481
 :||| :|||
 Db 714 KKRIGE FFFQVNYDWDLLAARSIWAFGPDSTGPNILVDDTL PSEVDKNLLTAVKDSIVQG 772

RESULT 5

AAM79209

ID AAM79209 standard; Protein; 858 AA.

XX

AC AAM79209;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human protein SEQ ID NO 1871.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US04098.

XX

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI; 2001-476283/51.

DR N-PSDB; AAK52342.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -

XX
PS Claim 20; Page 4256-4258; 6221pp; English.
XX

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX
SQ Sequence 858 AA;

Query Match 34.4%; Score 870; DB 22; Length 858;
Best Local Similarity 37.3%; Pred. No. 4.6e-75;
Matches 182; Conservative 95; Mismatches 189; Indels 22; Gaps 5;

QY 9 NLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTF-----GDI-----NYQEFARKLW 56
|::: |:|| | | : ||| |::| | : : |:||
Db 202 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261

QY 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGI 113
|| ||:| ||:| | : : |:| : ||:|:|: : : : |||
Db 262 GDRYFDPANGKFSKSATSPEGKKLPRTFCQLILDPIFKVFDALMNFKKEETAKLIEKLDI 321

QY 114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTTGGVDSDL 173
| | : : ||:| | : : : | | :||| | : | | | :
Db 322 KLDSEDKDKEGKPLLKAVMRRWLPA GDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 381

QY 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
: ||| |||| : : || | | : |||| || : | |::| ||| :||
Db 382 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKED 441

QY 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
: : | : : || : || || | : |||| :||| ||| | | | :|
Db 442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMVRMK 498

QY 294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVM 353
|: : |:::||| ||::|||:::|:|:| | | : :|||:| | ||:|: :
Db 499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIEESGEHIIAGAGELHLEICL 558

QY 354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
|| : : : | || :|||: : || | | : : | :||| |:| | | | ||||:|
Db 559 KDLEEDHACIPIKSDPVVSYRET VSEESNVLCLSKSPNKHNRLYMKARFPDGLAEDID 618

QY 414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
| | : : ||:|: || || ||| ||||| | | : |
Db 619 KGEVSARQELKQRARYLAEKYEWDVAEARKIWCFGPDGTGPNILTDITKGVQ----YLNE 674

QY 474 VKDSIVQG 481

Db :|||||
675 IKDSVVAG 682

RESULT 6

AAM80193

ID AAM80193 standard; Protein; 858 AA.

XX

AC AAM80193;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human protein SEQ ID NO 3839.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US04098.

XX

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI; 2001-476283/51.

DR N-PSDB; AAK53326.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX

PS Claim 20; Page 447; 6221pp; English.

XX

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX

SQ Sequence 858 AA;

Query Match 34.4%; Score 870; DB 22; Length 858;
Best Local Similarity 37.3%; Pred. No. 4.6e-75;
Matches 182; Conservative 95; Mismatches 189; Indels 22; Gaps 5;

```
QY      9 NLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTF-----GDI-----NYQEFARKLW 56
      |::: |:|| | | | : ||| ||::| | : : : |:||
Db     202 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261

QY     57 GDIYFNPKRKFTTKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
      || ||:| ||:| | : : |:| : ||:|::| : : : :| |
Db     262 GDRYFDPANGKFSKTSATSPGKKLPRTFCQLILDPIFKVFDALMNFKKEETAKLIEKLDI 321

QY     114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSL 173
      | |: :|||:|:::| : : | :||| | : | || | :
Db     322 KLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLSPVTAQKYRCELLYEGPPDDEA 381

QY     174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
      : ||| |||| : :|| | | :| : |||| | | : | |::| ||| :|
Db     382 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKED 441

QY     234 SQICTVGRWLWISVARYHIEVNRVPAGNWWVLIQVDPQPIVKTATITEPRGNEEAQIFRPLK 293
      : : | : : || : || || | : |||| :|| ||| | | | :|
Db     442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498

QY     294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVM 353
      |: : |:::| || ||::||:::|:::| || | : :|||::| | |||:| :
Db     499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICL 558

QY     354 HDLRKMYSEIDIKVADPVVTFCETVVETSSLKCFETPNKKNKITMIAEPLEKGLAEDIE 413
      || : : : | || :|||:: || | |:: | :::||| |:: | | | |||||:
Db     559 KDLEEDHACIPIKKSDFVVSRETVEESNVLCLSKSPNKHNRLYMKARFPDGLAEDID 618

QY     414 NEVVQITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
      | | : : |:|:| | || || ||| ||||| | | : |
Db     619 KGEVSARQELKQRARYLAKEYEWDVAEARKIWCFGPDGTGPNILTDITKGVQ----YLNE 674

QY     474 VKDSIVQG 481
      :|||:| |
Db     675 IKDSVVAG 682
```

RESULT 7

AAP70099

ID AAP70099 standard; Protein; 858 AA.

XX

AC AAP70099;

XX

DT 03-OCT-2002 (updated)
DT 18-APR-1991 (first entry)
XX
DE Sequence of elongation factor 2 (EF-2) in plasmid pHEW 1.
XX
KW Protein biosynthesis; diphtheria toxin; Ps aeruginosa toxin;
KW resistant; selective marker.
XX
OS Rattus rattus.
XX
PN EP217327-A.
XX
PD 08-APR-1987.
XX
PF 29-SEP-1986; 86EP-0113349.
XX
PR 30-SEP-1985; 85JP-0219545.
XX
PA (SUMO) SUMITOMO CHEM IND KK.
XX
PI Uchida T;
XX
DR WPI; 1987-095363/14.
DR N-PSDB; AAN70159.
XX
PT New DNA sequence encoding elongation factor 2 - is useful as
PT selective marker for animal cells transformed with exogenous DNA
PT from un-transformed cells
XX
PS Disclosure; Fig. 1; 38pp; English.
XX
CC EF-2 is esssntial for the translocation of peptidyl- tRNA from A-
CC site to P-site on the ribosome. It is inactivated through formation
CC of a covalent bond with ADP ribose derived from NAD by the action of
CC diphtheria toxin or Ps. aeruginosa exotoxin, and so cells exposed to
CC the toxin may die. A variant EF-2 (claimed) is resistant to diphtheria
CC toxin and Pseudomonas aeruginosa exotoxin.
CC (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ Sequence 858 AA;

Query Match 34.2%; Score 863; DB 8; Length 858;
Best Local Similarity 37.1%; Pred. No. 2.2e-74;
Matches 181; Conservative 95; Mismatches 190; Indels 22; Gaps 5;

QY 9 NLILSPLLGNVCFSSSQYSICFTLGSAFIYADTF-----GDI-----NYQEFARLW 56
|::: |::| | | : | | | |::| | : :: |::|
Db 202 NIMIDPVLGTVGFGSGLHGWFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261
QY 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
| | |::| |::| | : : |::| |::|::|: :: :::| |
Db 262 GDRYFDPANGKFSKANS PDGKKLPRTFCQLILDPIFKVFDAMNFRKEETAKLIEKLDI 321
QY 114 HLTKEELKLNIRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
| | : : |::| |::: : | |::| | : | | | :
Db 322 KLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLPSVTAQKYRCELLYEGPPDDEA 381

Qy 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEED 233
 : ||| |||| : :|| | | :|: |||| | | : | |::| ||| :|:
 Db 382 AMGIKSCDPKGPMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKEE 441
 Qy 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
 : : | : : || : || || | : |||| :||| ||| | | | :|
 Db 442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498
 Qy 294 FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVM 353
 |: : |::||| ||::|||::||::| | | : :|||::| | ||::|: :
 Db 499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICL 558
 Qy 354 HDLRKMYSEIDIKVADPVVTFCETVVETSSLKCFEAETPNKKNKITMIAEPLKGLAEDIE 413
 || : : | || :|||:: ||| | |:: | :::||| |:: | | | |||||:
 Db 559 KDLEEDHACIPIKSDPVVSYRETVSEESNVLCLSKSPNKHNRLYMKARFPDGLAEDID 618
 Qy 414 NEVVQITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
 | | : | : ||: ||| : || || |||| ||||| | | : --- |
 Db 619 KGEVSARQELKARARYLAEKYEWDAEARKIWCFGPDGTGPNILTDITKGVQ---YLNE 674
 Qy 474 VKDSIVQG 481
 :|||:| |
 Db 675 IKDSVVAG 682

RESULT 8

AAP70100

ID AAP70100 standard; Protein; 858 AA.

XX

AC AAP70100;

XX

DT 03-OCT-2002 (updated)

DT 18-APR-1991 (first entry)

XX

DE Sequence of variant of elongation factor 2 (EF-2), in
 DE plasmid pHEd 1, resistant to diphtheria toxin and Pseudomonas
 DE aeruginosa exotoxin.

XX

KW Protein biosynthesis; diphtheria toxin; Ps aeruginosa toxin;
 KW resistant; selective marker.

XX

OS Rattus rattus.

XX

PN EP217327-A.

XX

PD 08-APR-1987.

XX

PF 29-SEP-1986; 86EP-0113349.

XX

PR 30-SEP-1985; 85JP-0219545.

XX

PA (SUMO) SUMITOMO CHEM IND KK.

XX

PI Uchida T;

XX

DR WPI; 1987-095363/14.

DR N-PSDB; AAN70160.

XX
PT New DNA sequence encoding elongation factor 2 - is useful as
PT selective marker for animal cells transformed with exogenous DNA
PT from un-transformed cells
XX
PS Disclosure; Fig 2; 38pp; English.
XX
CC EF-2 is esssntial for the translocation of peptidyl- tRNA from A-
CC site to P-site on the ribosome. It is inactivated through formation
CC of a covalent bond with ADP ribose derived from NAD by the action of
CC diptheria toxin or Ps. aeruginosa exotoxin, and so cells exposed to
CC the toxin may die. A variant EF-2 (claimed) is resistant to diptheria
CC toxin and Pseudomonas aeruginosa exotoxin.
CC (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ Sequence 858 AA;

Query Match 34.2%; Score 863; DB 8; Length 858;
Best Local Similarity 37.1%; Pred. No. 2.2e-74;
Matches 181; Conservative 95; Mismatches 190; Indels 22; Gaps 5;

Qy 9 NLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTF-----GDI-----NYQEFARKLW 56
|::: |:| | | : || |::| | : :: |:|
Db 202 NIMIDPVLGTVGFGSLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261
Qy 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
|| |:| |:| | : : |:| : |:|::|: :: :::| |
Db 262 GDRYFDPANGKFSKSANSPDGKKLPRTFCQLILDPIFKVFDAIMNFRKEETAKLIEKLDI 321
Qy 114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGA PKIEHTYTGGVDS DL 173
| |: :||: |::: : | :|| | : | | :
Db 322 KLDSEDKDKEGKPLLKAVMRWLPA GDALLQMITIHLSPVTAQKYRCELLYEGPPDDEA 381
Qy 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
: ||| ||| : :|| | | :||| || : | :::| ||| :|:
Db 382 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVSTGLKVRIMGPNYTPGKKEE 441
Qy 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
: : | : : || : || || | : ||| :||| ||| | | | :|
Db 442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498
Qy 294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYP SLTKVEESGEHVILGTGELYLDCVM 353
|: : |::||| ||::||::||::||: || | : :|||::| | ||::|: :
Db 499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIAGAGELHLEICL 558
Qy 354 HDLRKMYSEIDIKVADPVVTFCE TVVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
|| : : | || :|||:: || | :| :|::||| :|: | | | |||||:
Db 559 KDLEEDHACIPIKSDPVVS YRET VSEES NVLCLSKSPNKHNRLYMKARFPDGLAEDID 618
Qy 414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
| | : ||::|: || || ||| ||||| | | : |
Db 619 KGEVSARQELKARARYLAKEYEWDVAEARKIWCFGPDGTGPNILTDITKGVQ----YLNE 674
Qy 474 VKDSIVQG 481
:|::| |
Db 675 IKDSVVAG 682

RESULT 9

ABB67414

ID ABB67414 standard; Protein; 699 AA.

XX

AC ABB67414;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 29034.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR N-PSDB; ABL11517.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX

PS Disclosure; SEQ ID NO 29034; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 699 AA;

Query Match 33.4%; Score 844.5; DB 22; Length 699;

Best Local Similarity 37.2%; Pred. No. 9.9e-73;

Matches 175; Conservative 103; Mismatches 183; Indels 9; Gaps 4;

QY 12 LSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFARLWGDIYFNPKTRKFTKK 71

```

      : | | : | | : | | | | : : : | | : : | | | : |
Db    196 VDPSKGSVGFSGSLHGWAFTLKQFSEMYSEKF-KIDVVKLMNRLWGENFFNAKTKKW-QK 253

      : : : | | : | : | : : : : : | : : : | : |
Qy    72 APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLLRLV 131
      : : | | : | : | : : : : : | : : : | : |
Db    254 QKEADNKRSFCMYILDPIYKVFDAIMNYKKEEIGTLLEKIGVTLKHEDKDKDGKALLKTV 313

      : : : | | : | : | : : : | | : | : | | | : :
Qy    132 CKKFFGEFTGFVDMCVQHIPSQVGAQPKIEHTYTTGGVSDSLGEAMSDCDPDGGLMCHTT 191
      : : : | | : | : | : : : | | : | : | | | : :
Db    314 MRTWLPAGEALLQMIATHLPSPVVAQKYRMEMLYEGPHDDEAAIAVKSCDPDGGLMMYIS 373

      : : | | : | : | : : : | | : | : | : | : : | |
Qy    192 KMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYHI 251
      : : | | : | : | : : : | | : | : | : | : : | |
Db    374 KMVPTSDKGRFYAFGRVFAGKVATGQKCRIMGPNYTPGKKEDLYEKAIQRTILMMGRYVE 433

      : | : | : | : | : | : : : : : : : | : : | : :
Qy    252 EVNRVPAGNWVLIEGVDQPIVKTTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE 311
      : | : | : | : | : | : : : : : : : | : : | : :
Db    434 AIEDVPSGNICGLVGVDQFLVKGTGTITT---FKDAHNMKVMKFSVSPVVRVAVEPKNPAD 490

      : : : | : : | : : | : : | : : | : : | : : | : :
Qy    312 LPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPV 371
      : : : | : : | : : | : : | : : | : : | : : | : :
Db    491 LPKLVEGLKRLAKSDPMVQCIIEESGEHI IAGAGELHLEICLKDLEEDHACIPLKKS DPV 550

      : : | | : | : | : : : | | : | : | : | : : | |
Qy    372 VTFCETVETSSSLKCFATPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFFQ 431
      : : | | : | : | : : : | | : | : | : | : : | |
Db    551 VSYRETVSEESDQMCLSKSPNKHNRLLMKALPMPDGLPEDIDNGDVSADKDEFKARARYLS 610

      : : | | : | : | : : : | | : | : | : | : : | |
Qy    432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
      : : | | : | : | : : : | | : | : | : | : : | |
Db    611 EKYDYDVTEARKIWCFGPDGTGPNFILDCTKSVQ----YLNEIKDSVVAG 656

```

RESULT 10

ABB58973

ID ABB58973 standard; Protein; 832 AA.

XX

AC ABB58973;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 3711.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL03076.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 3711; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 832 AA;

Query Match 33.4%; Score 844.5; DB 22; Length 832;
 Best Local Similarity 37.2%; Pred. No. 1.3e-72;
 Matches 175; Conservative 103; Mismatches 183; Indels 9; Gaps 4;

QY 12 LSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDYFNPKTRKFTKK 71
 : | | : | | : ||| |:::| : | : ||| : || | : |
 Db 196 VDPSKGSVGFSGSLHGWAFTLKQFSEMYSEKF-KIDVVKLMNRLWGENFFNAKTKKW-QK 253
 QY 72 APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLNIRPLRLV 131
 : ::||| : ||:| : : : |:::| | : : : ||: |
 Db 254 QKEADNKRSCFMYILDPIYKVFDAIMNYKKEEIGTLLEKIGVTLKHEDKDKDGKALLKTV 313
 QY 132 CKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGVDSDLGEAMSDCDPDG PLMCHTT 191
 : : : | | : ||| | | : | | : | : ||||| : :
 Db 314 MRTWLPAGEALLQMI AIHLPS PVVAQKYRMEMLYEGPHDDEAAI AVKSCDPDG PLMMYIS 373
 QY 192 KMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYHI 251
 || | | :|: ||| :| : || :::| ||| :|| : | : : ||
 Db 374 KMOVPTSDKGRFYAFGRVFAGKVATGQKCRIMGPNYTPGKKEDLYEKAIQRTILMMGRYVE 433
 QY 252 EVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE 311
 : ||:| | : ||| :||| ||| : : | : :| : :|::||| ||::
 Db 434 AIEDVPSGNICGLVGVDQFLVKTGTITT---FKDAHNMKVMKFSVSPVVRVAVEPKNPAD 490
 QY 312 LPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYS EIDIKVADPV 371
 |||::|::| : || : :||| :| | |||: : | | : : | :| :||
 Db 491 LPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPLKKS DPV 550
 QY 372 VTFCE TVVETSSLKCAETPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRKKLG EFFQ 431
 |:: ||| | | : :||| |:: | | | : | | |||: | | :
 Db 551 VSYRETVSEESDQMCLSKSPNKHNRLLMKALPMPDGLPEDIDNGDVS AKDEFKARARYLS 610

Qy 432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSSEVDKALLGSVKDSIVQG 481
 |||:|: || || |||| |||| ::| | : | :|||:| |
 Db 611 EKYDYDVTEARKIWCFGPDGTGPNFILDCTKSVQ---YLNEIKDSVVAG 656

RESULT 11

AA43637

ID AAY43637 standard; Protein; 842 AA.

XX

AC AAY43637;

XX

DT 11-FEB-2000 (first entry)

XX

DE Amino acid sequence of a translation elongation factor-2 (EF-2).

XX

KW Translation elongation factor-2; EF-2; yeast; diphtheria; toxin;
 KW diphthamide; DPH1; DPH2; DPH3; DPH4; DPH5; toxin resistance; EFT1;
 KW selective killing; microorganism; microcidal toxin; biological control.

XX

OS Saccharomyces cerevisiae.

XX

FH Key Location/Qualifiers

FT Modified-site 699

FT /note= "this residue is post-translationally
 FT modified to diphthamide by the DPH gene
 FT products"

XX

PN WO9953762-A1.

XX

PD 28-OCT-1999.

XX

PF 16-APR-1999; 99WO-US08213.

XX

PR 17-APR-1998; 98US-0082089.

XX

PA (MICR-) MICROBIA INC.

XX

PI Milne GT, Fink GR;

XX

DR WPI; 2000-013163/01.

DR

N-PSDB; AAZ30303.

XX

PT Selective biological control of microorganisms -

XX

PS Example 2; Fig 3A-E; 75pp; English.

XX

CC The present sequence represents the translation elongation factor-2
 CC (EF-2), designated EFT1 (for EFT2 see AAY43638). Diphtheria toxin
 CC catalyses the ADP-ribosylation of EF-2 on a diphthamide residue. Mutants
 CC that fail to make diphthamide or lack the histidine precursor of
 CC diphthamide are resistant to the toxin. Therefore, mutations in genes
 CC involved in this (EF-2 (2 genes), DPH1, DPH2, DPH3, DPH4, and DPH5) are
 CC sufficient to provide resistance to diphtheria toxin. Deletion of any of
 CC the DPH genes is sufficient to yield nearly complete resistance to
 CC diphtheria toxin. The DPH toxin phenotype is completely recessive,
 CC while the EF-2 toxin phenotype is generally dominant. This is utilised

CC in the method of the invention. The specification describes a method
 CC for the selective killing of a microorganism (M1). This is achieved by
 CC contacting it with another microorganism (M2) producing a microcidal
 CC toxin such as diphtheria. M1 and M2 are allowed to fuse and thus the
 CC toxin is delivered into, and kills the microorganism formed following
 CC fusion. M2 has been modified so that it is resistant to the toxin,
 CC e.g. by introducing mutations in DPH1, DPH3 or DPH4 genes that decreases
 CC biosynthesis of diphthamide. The method is used for selective biological
 CC control of harmful microorganisms.

XX

SQ Sequence 842 AA;

Query Match 32.0%; Score 807.5; DB 21; Length 842;
 Best Local Similarity 35.7%; Pred. No. 5.4e-69;
 Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;

Qy 14 PLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFARLWGDYFNPKTRKFTKK-- 71
 | | | | : ||: || || || :: : |||| :|||:|:| |
 Db 204 PARGTVAFGSLHGWAFTRQFATRYAKKFG-VDKAKMMDRLWGDSEFNPKTKKWTNKDT 262
 Qy 72 -APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLNIRPLRL 130
 | :|| |||:|::: : : ||:| | | : | :||:
 Db 263 DAEGKPLERAENMFILDPFRLFTAIMNFKKDEIPVLEKLEIVLKGDEKDLEGKALLKV 322
 Qy 131 VCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHT 190
 | :|| :|| | | || : | | | | : :||| || :
 Db 323 VMRKFLPAADALLEMIVLHLPSPVTAQAYRAEQLYEGPADDANCIAIKNCDPKADLMYLV 382
 Qy 191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYH 250
 :|| | | :|: |||| :||: || |:: | || :|| | :|: :|:
 Db 383 SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPGKKDDLFKAIQRVVLMMGRFV 442
 Qy 251 IEVNRVPAGNWVLIEGVDQPIVKATITTEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
 :: |||| : : |:|| :|| |:: :| | :||: : |::||| | :
 Db 443 EPIDDCPAGNIIGLVGIDQFLKLTGTLTT---SETAHNMKVMKFSVSPVVQVAVEVKAN 499
 Qy 311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADP 370
 :|||::|:|:|:| | : | : ||||: ||||:|: : || : : :|:| |
 Db 500 DLPKLVEGLKRLSKSDPCVLTYMSESGEHIVAGTGELHLEICLDLEHDHAGVPLKISPP 559
 Qy 371 VVTFCEVTVETSSLKCFATPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFF 430
 || : ||| || :||| |:: |||:|: : ||| : : : |
 Db 560 VVAYRETVESESSQTALSKSPNKHNRILYKAEPIDEEVSLAIENGIINPRDDFKARARIM 619
 Qy 431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
 | ||: || || |||| |||:| | ||: | :|||:
 Db 620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSV 664

RESULT 12

AA43638

ID AA43638 standard; Protein; 842 AA.

XX

AC AA43638;

XX

DT 11-FEB-2000 (first entry)

XX

DE Amino acid sequence of a translation elongation factor-2 (EF-2).
XX
KW Translation elongation factor-2; EF-2; yeast; diphtheria; toxin;
KW diphthamide; DPH1; DPH2; DPH3; DPH4; DPH5; toxin resistance; EFT1;
KW selective killing; microorganism; microcidal toxin; biological control.
XX
OS *Saccharomyces cerevisiae*.
XX
FH Key Location/Qualifiers
FT Modified-site 699
FT /note= "this residue is post-translationally
FT modified to diphthamide by the DPH gene
FT products"
XX
PN WO9953762-A1.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US08213.
XX
PR 17-APR-1998; 98US-0082089.
XX
PA (MICR-) MICROBIA INC.
XX
PI Milne GT, Fink GR;
XX
DR WPI; 2000-013163/01.
DR N-PSDB; AAZ30304.
XX
PT Selective biological control of microorganisms -
XX
PS Example 2; Fig 4A-E; 75pp; English.
XX
CC The present sequence represents the translation elongation factor-2
CC (EF-2), designated EFT2 (for EFT1 see AAY43637). Diphtheria toxin
CC catalyses the ADP-ribosylation of EF-2 on a diphthamide residue. Mutants
CC that fail to make diphthamide or lack the histidine precursor of
CC diphthamide are resistant to the toxin. Therefore, mutations in genes
CC involved in this (EF-2 (2 genes), DPH1, DPH2, DPH3, DPH4, and DPH5) are
CC sufficient to provide resistance to diphtheria toxin. Deletion of any of
CC the DPH genes is sufficient to yield nearly complete resistance to
CC diphtheria toxin. The DPH toxin phenotype is completely recessive,
CC while the EF-2 toxin phenotype is generally dominant. This is utilised
CC in the method of the invention. The specification describes a method
CC for the selective killing of a microorganism (M1). This is achieved by
CC contacting it with another microorganism (M2) producing a microcidal
CC toxin such as diphtheria. M1 and M2 are allowed to fuse and thus the
CC toxin is delivered into, and kills the microorganism formed following
CC fusion. M2 has been modified so that it is resistant to the toxin,
CC e.g. by introducing mutations in DPH1, DPH3 or DPH4 genes that decreases
CC biosynthesis of diphthamide. The method is used for selective biological
CC control of harmful microorganisms.
XX
SQ Sequence 842 AA;

Query Match 32.0%; Score 807.5; DB 21; Length 842;
Best Local Similarity 35.7%; Pred. No. 5.4e-69;

Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;

```

QY      14 PLLGNVCFSSSQYSICFTLGSEAKIYADTFGDINYQEFAKRLWGDYFNPKTRKFTKK-- 71
      | | | | | : | | | | | : : | | | | : | | | | : | |
Db     204 PARGTVAFGSGLHGWAFTRQFATRYAKKFG-VDKAKMMDRLWGDSFFNPKTKKWTNKDT 262

QY      72 -APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL 130
      | : | | | | | : | | : | | : | | : | | : | | : | |
Db     263 DAEGKPLERAFNMFILDPIFRLFTAIMNFKKDEIPVLLLEKLEIVLKGDEKDLEGKALLKV 322

QY     131 VCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPIMCHT 190
      | : | | | | : | | | | | : | | | | : | | | | : | |
Db     323 VMRKFLPAADALLEMIVLHLPSPVTAQAYRAEQLYEGPADDANCIAIKNCDPKADLMLYV 382

QY     191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYLTLEDEEDSQICTVGRLWISVARYH 250
      : | | | | : | | | | : | | : | | : | | : | | : | |
Db     383 SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPGKKDDLFIKAIQRRVVLMMGREV 442

QY     251 IEVNRVPAGNWWLIEGVQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
      : : | | | : : | | : | | : | | : | | : | | : | | : | |
Db     443 EPIDDCPAGNIIGLVGIDQFLLKTGTLT---SETAHNMKVMKFSVSPVVQVAVEVKAN 499

QY     311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSIEDIKVADP 370
      : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db     500 DLPKLVEGLKRLSKSDPCVLTYMESGEHIVAGTGELHLEICLQDLEHDHAGVPLKISPP 559

QY     371 VVTFCETVVETSSSLKCFATPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRKKLGEFF 430
      | | : | | | | : | | : | | : | | : | | : | | : | |
Db     560 VVAYRETVESESSQTALSKSPNKHNRILKAEPIDEEVSLAIENGIINPRDDFKARARIM 619

QY     431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
      | | | : | | | | | | | | | | | | | | | | | | | | | |
Db     620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSVV 664

```

RESULT 13

ABJ25942

ID ABJ25942 standard; Protein; 884 AA.

XX

AC ABJ25942;

XX

DT 16-APR-2003 (first entry)

XX

DE Aspergillus fumigatus essential gene protein #600.

XX

KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response.

XX

OS Aspergillus fumigatus.

XX

PN WO200286090-A2.

XX

PD 31-OCT-2002.

XX

PF 23-APR-2002; 2002WO-US13142.

XX

PR 23-APR-2001; 2001US-285697P.

PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX

DR WPI; 2003-093124/08.

XX

PT New purified or isolated nucleic acids of essential genes of
PT *Aspergillus fumigatus*, useful for treating or preventing infections by
PT *A. fumigatus*, or for treating a non-infectious disease in a subject
PT e.g. cancer -

XX

PS Disclosure; Page -; 175pp; English.

XX

CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of *A.*
CC *fumigatus* to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This sequence represents a protein of one of the essential genes
CC of *Aspergillus fumigatus* of the invention.

XX

SQ Sequence 884 AA;

Query Match 31.9%; Score 806; DB 24; Length 884;

Best Local Similarity 36.0%; Pred. No. 8.3e-69;

Matches 169; Conservative 104; Mismatches 183; Indels 14; Gaps 5;

Qy 14 PLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFARLWGDYFNPKTRKFTKKAP 73
| | | | | : ||: || || || :: :: :||| | |||:|:|:
Db 201 PDRGTVAFGSGLHGWAFTVRQFAVKYAKKFG-VDRKKMLERLWGDNYFNPQTKKWTGTG- 258
Qy 74 TSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKLEELKLNIRPLRLVCK 133
||:| |||:|:|:| | | : : ::| | | :| | : ||:::

Db 259 --EPEQRAFNMFILDPIFKIFAAVNNDKTEEIHKLVEKLEIKLASDEKDLKGKALLKVIMR 317

QY 134 KFFGEFTGFVDMCVQHIPSQVGAQPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHTTKM 193
 || :| |:|| | :| | :| : ||| ||| :||

Db 318 KFLPAADAMLEMICIHLPSPVTAQKYRAETLYEGPMDDECAIGIRDCDPKAPLMLYVSKM 377

QY 194 FSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEDEEDSQICTVGRLWISVARYHIEV 253
 | | :| :||| || :| | :| | :| :| :| :| :

Db 378 VPTSDKGRFYAFGRVFSGIVKSGLVRIQGPNIYIPGKKDDL FVKAIQRTILMMGREVEPI 437

QY 254 NRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSELP 313
 |||| | : |||| :| :| :| | :| :| :| :| :| :

Db 438 EDVPAGNIVGLVGVDQFLLKSGTLTT---SETAHLNLMKFSVSPVVRGVEVKNAQDLP 494

QY 314 KMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRMYSEIDIKVADPVVT 373
 |::||::|| | : | : |||:|: | |||:| : || : : : :|||:

Db 495 KLVEGLKRLSKSDPCVLTMINESGQHIVAGAGELHLEICLDLEEDHAGVPLRISDPVVS 554

QY 374 FCETVETSSSLKCFEAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGFEFFQTK 433
 : ||| ||: :||| |:: : |:: : : : ||: : : |

Db 555 YRETVGGESSMTALSKSPNKHNRLYVTAQPLGEEVSLATESGKINPRDDFKARARLLADD 614

QY 434 YDWDLAARSIWAFGPDATGPNILVDDTL PSEVDKAL--LGSVKDSIVQG 481
 | ||: || || |||| || |:||| | ||: | :||| | |

Db 615 YGWDVTDARKIWCFGPDTTGANLLVDQT-----KAVQYLNEIKDSFVSG 658

RESULT 14

AAB71315

ID AAB71315 standard; Protein; 845 AA.

XX

AC AAB71315;

XX

DT 18-NOV-2002 (first entry)

XX

DE L. major 4G2-83 extended antigen SEQ ID 110.

XX

KW Antigen; immunogenic; antiparasitic; immunostimulant; leishmaniasis;
 KW gene therapy; vaccine; interleukin-12 agonist.

XX

OS Leishmania major.

XX

PN US2002081320-A1.

XX

PD 27-JUN-2002.

XX

PF 04-JUN-2001; 2001US-0874923.

XX

PR 22-SEP-1995; 95US-0533669.

PR 12-FEB-1997; 97US-0798841.

PR 27-AUG-1997; 97US-0920609.

PR 12-FEB-1998; 98US-0022765.

PR 30-OCT-1998; 98US-0183861.

PR 14-APR-2000; 2000US-0551974.

PR 05-MAY-2000; 2000US-0565501.

PR 14-AUG-2000; 2000US-0639206.

XX

PA (REED/) REED S G.
 PA (CAMP/) CAMPOS-NETO A.
 PA (WEBB/) WEBB J R.
 PA (DILL/) DILLON D C.
 PA (SKEI/) SKEIKY Y A W.
 PA (BHAT/) BHATIA A.
 PA (COLE/) COLER R N.
 PA (PROB/) PROBST P.
 PA (BRAN/) BRANNON M.

XX

PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;
 PI Coler RN, Probst P, Brannon M;

XX

DR WPI; 2002-635457/68.

DR N-PSDB; AAF88573.

XX

PT New polypeptide containing at least an immunogenic portion of one or
 PT more Leishmania antigens or their variants, useful for preventing,
 PT treating and detecting leishmaniasis, and stimulating immune responses
 PT in patients -

XX

PS Claim 1; Page 108-110; 163pp; English.

XX

CC This invention describes a novel polypeptide containing an immunogenic
 CC portion of a Leishmania antigen or its variant which has antiparasitic
 CC and immunostimulant activity. The compositions and methods of the present
 CC invention are useful for preventing, treating and detecting
 CC leishmaniasis, and stimulating immune responses in patients against
 CC leishmaniasis. The polypeptides and the polynucleotides encoding them can
 CC be used for gene therapy, in vaccines or as interleukin-12 agonists. The
 CC compositions and methods of the present invention, as compared to prior
 CC art, are more improved therapeutic modalities in the diagnosis,
 CC prevention and treatment of leishmaniasis.

XX

SQ Sequence 845 AA;

Query Match 29.4%; Score 743; DB 23; Length 845;

Best Local Similarity 34.7%; Pred. No. 1.1e-62;

Matches 166; Conservative 102; Mismatches 189; Indels 22; Gaps 9;

Qy 12 LSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFARLWGDIFYNPKTRKFTKK 71
 :|| | | : || || :|| : : :||| :| :| :|
 Db 200 VSPEKGTVAIGSGLQAWAFSLTRFANMYAAKFG-VDELKMRERLWGDNFFDAKNKKWIKQ 258
 Qy 72 APTSSSQ---RSFVEFILEPLYKILAQVVGDDVDTSLPRTLDELGIHLTKEELKLNIRPLL 128
 : : || :| ||| :| : : :| :|| || : ||
 Db 259 ETNADGERVRRAFQCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQVPXKLL 318
 Qy 129 RLVCKKFFGEFTGFVDMCVQHIPS PKVGA KP KIEHTYTGGVDS D----LGEAMSDCDPDG 184
 : | | : | | ||| : | || : :| :|||
 Db 319 KTVMMXFLPAAETLLQMIVAHLPSPKKAQAYRAEMLYSGEASPEDKYFMG--IKNCDPAA 376
 Qy 185 PLMCHTTKMFSTHDGVQHFHFGRLVSGTIHAGQPVKVLGENYTLDEED--SQICTVGRWLW 243
 ||| : || | | :| ||| :| :||| :|| :| :|
 Db 377 PLMLYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDKPVQRSV 436
 Qy 244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303

```

      : : || | :| || | : |||: |||:||||: | | :|:: : |::|
Db      437 LMMGRYQEAVEDMPCGNVVGVLGVVDKYIVKSATITD--DGESPHPLRDMKYSVSPVVRVA 494
QY      304 VEPVNPSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMY-SE 362
      || |||:||||::|::: || | : :|||||::| |||:|: : ||:: : :
Db      495 VEAKNPSDLPKLVEGLKRLAKSDPLVVCSIEESGEHIVAGAGELHLEICLKDLQEDFMNG 554
QY      363 IDIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAEPLEKGLAEDIENEVVQITWN 422
      :|:::||||:| ||| : || :| ::| || |:: || : || :|
Db      555 APLKISEPVVSFRET VTDVSSQQCLS KSKANKHNRLFCRGAPLTEELALAMEEGTAGPEAD 614
QY      423 RKKLGEEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKAL--LGSVKDSIV 479
      | | | :||: || || :||| |||::|| | | : : :||| |
Db      615 PKVRARFLADNYEWDVQEARKIW CYGPDNRGPNVVVD-----VTKGVQNMAEMKDSFV 667

```

RESULT 15

AAU71861

ID AAU71861 standard; Protein; 845 AA.

XX

AC AAU71861;

XX

DT 26-FEB-2002 (first entry)

XX

DE Leishmania antigen 4G2-83 extended protein.

XX

KW Leishmaniasis; Leishmania antigen; immunostimulant; protozoacide; Ldp23;

KW interleukin-15; Lbhsp83; M15; Lt-1; LbeIF4A; Lmspla; Lmsp9a; MAPS-1A;

KW LmgSP1; LmgSP3; LmgSP5; LmgSP8; LmgSP9; LmgSP13; LmgSP19; LcgSP1; LcgSP3;

KW LcgSP4; LcgSP8; LcgSP10; 1G6-34; 1E6-44; 4A5-63; 1B11-39; 2A10-37;

KW 4G2-83; 4H6-41; 8G3-100.

XX

OS Leishmania major.

XX

PN WO200179276-A2.

XX

PD 25-OCT-2001.

XX

PF 05-APR-2001; 2001WO-US11254.

XX

PR 14-APR-2000; 2000US-0551974.

PR 05-MAY-2000; 2000US-0565501.

PR 14-AUG-2000; 2000US-0639206.

XX

PA (CORI-) CORIXA CORP.

XX

PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;

PI Coler RM, Probst P;

XX

DR WPI; 2002-061971/08.

DR N-PSDB; AAS96071.

XX

PT New isolated Leishmania antigens, useful for prevention, treatment and

PT diagnosis of leishmaniasis, also related nucleic acids for genetic

PT vaccination -

XX

PS Claim 1; Page 190-192; 193pp; English.

XX
 CC The invention relates to polypeptides comprising an immunogenic part of a
 CC Leishmania antigen. The Leishmania polypeptides and their associated DNA
 CC sequences, epitopes and fusion proteins are used in the production of
 CC compositions used for inducing a protective immune response against
 CC leishmaniasis, for prevention and treatment of the disease. The
 CC compositions can also be used generally to treat diseases that respond to
 CC interleukin-15 stimulation. In addition, the products may contain an
 CC immunostimulant. Sequences AAU71805-AAU71862 represent Leishmania
 CC antigens and antigenic peptides of the invention.
 XX
 SQ Sequence 845 AA;

Query Match 29.4%; Score 743; DB 23; Length 845;
 Best Local Similarity 34.7%; Pred. No. 1.1e-62;
 Matches 166; Conservative 102; Mismatches 189; Indels 22; Gaps 9;

Qy 12 LSPLLGNVCFSSSQYSICFTLGSAFKIYADTFGDINYQEFKRLWGDYFNPKTRKFTKK 71
 :|| | | : | :| || :|| || : : : ||||| :| :| :| :|
 Db 200 VSPEKGTVAIGSGLQAWAFSLTRFANMYAAKFG-VDELMKRERLWGDNFFDAKNKKWIKQ 258
 Qy 72 APTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLL 128
 : : | :| :| :| :| :| : : : | | : || || : ||
 Db 259 ETNADGERVRR AFCQFC LDP IYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQVPXKLL 318
 Qy 129 RLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSD---LGEAMSDCDPDG 184
 : | | : | | | :||| : | | :| : :| : :|||
 Db 319 KTVMMXFLPAAETLLQMI VAHLPSPKKAQAYRAEMLYSGEASPEDKYFMG--IKNCDPAA 376
 Qy 185 PLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEED-SQICTVGRWLW 243
 ||| : :|| | | :| ||| : || : :|| | || :|| : | |
 Db 377 PLMLYISKMVPTADRGRFFAEGRI FSGKVRSGQKVRIMGN NYVYGKKQDLYEDKPVQRSV 436
 Qy 244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
 : : || | :| || | : ||| : ||| :||| : | | :| : : :| : :|
 Db 437 LMMGRYQEAVEDMPCGNVVGVLGVGD KYIVKSATITD--DGESPHPLRDMKYSVSPVVRVA 494
 Qy 304 VEPVNPSEL PKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLRKMY-SE 362
 || ||| :||| : :||| : || : :||| :||| : ||| :| : ||| : : :
 Db 495 VEAKNP SDLPKLVEGLKRLAKSDPLVVC SIEESGEHIVAGAGELHLEICLKD LQEDFMNG 554
 Qy 363 IDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAEPLEKGLAEDIENEV VQITWN 422
 :||| :||| : ||| : || :| : : || :| : || : || :| :
 Db 555 APLKISEPVVSFRET VTDVSSQQCLSKSANKHNRLFCRGAPLTEELALAMEEGTAGPEAD 614
 Qy 423 RKKLGEFFQTKYDWDLLAARS IWAFGPDATGPNILVDDTL PSEVDKAL--LGSVKDSIV 479
 | | | :||| : || | :||| ||| :||| | | : : :||| |
 Db 615 PKVRARFLADNYEWDVQEARKIW CYGPDN RGNPVVVD-----VTKG VQNMAEMKDSFV 667

Search completed: January 30, 2004, 11:24:34
 Job time : 73.1206 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 11:23:12 ; Search time 27.1381 Seconds
(without alignments)
749.923 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSVKDSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2527	100.0	481	4	US-08-928-862-3	Sequence 3, Appli
2	807.5	32.0	842	4	US-09-293-549-2	Sequence 2, Appli
3	807.5	32.0	842	4	US-09-293-549-4	Sequence 4, Appli
4	554.5	21.9	361	4	US-09-183-861-85	Sequence 85, Appl
5	554.5	21.9	361	4	US-09-022-765-85	Sequence 85, Appl
6	554.5	21.9	361	4	US-09-551-974A-85	Sequence 85, Appl
7	179	7.1	900	4	US-09-252-991A-25011	Sequence 25011, A
8	177	7.0	716	4	US-09-252-991A-21345	Sequence 21345, A
9	175	6.9	705	4	US-09-134-001C-5356	Sequence 5356, Ap
10	148.5	5.9	651	4	US-09-198-452A-589	Sequence 589, App
11	113.5	4.5	277	4	US-09-328-352-4749	Sequence 4749, Ap

12	111	4.4	1067	4	US-09-252-991A-19697	Sequence 19697, A
13	105	4.2	1224	4	US-09-107-532A-6220	Sequence 6220, Ap
14	103	4.1	1380	4	US-09-328-352-8132	Sequence 8132, Ap
15	100.5	4.0	677	4	US-09-252-991A-32924	Sequence 32924, A
16	97.5	3.9	616	4	US-09-328-352-5348	Sequence 5348, Ap
17	95	3.8	738	6	5264554-2	Patent No. 5264554
18	94	3.7	828	1	US-08-261-304-2	Sequence 2, Appli
19	93.5	3.7	995	5	PCT-US95-04910-14	Sequence 14, Appl
20	93	3.7	738	3	US-08-478-208-32	Sequence 32, Appl
21	93	3.7	738	4	US-09-336-536-73	Sequence 73, Appl
22	92.5	3.7	531	4	US-09-134-001C-3574	Sequence 3574, Ap
23	92	3.6	479	4	US-09-134-001C-4128	Sequence 4128, Ap
24	92	3.6	732	4	US-08-671-757A-7	Sequence 7, Appli
25	92	3.6	732	4	US-08-671-757A-8	Sequence 8, Appli
26	91	3.6	240	4	US-09-107-532A-4175	Sequence 4175, Ap
27	91	3.6	425	4	US-09-134-001C-2895	Sequence 2895, Ap
28	90.5	3.6	527	3	US-08-369-822C-25	Sequence 25, Appl
29	90.5	3.6	527	3	US-08-582-776C-40	Sequence 40, Appl
30	90.5	3.6	527	3	US-08-434-831B-37	Sequence 37, Appl
31	90	3.6	499	5	PCT-US96-03916-8	Sequence 8, Appli
32	90	3.6	499	5	PCT-US96-03916-69	Sequence 69, Appl
33	89.5	3.5	634	4	US-09-134-001C-3390	Sequence 3390, Ap
34	89.5	3.5	840	2	US-08-467-822-25	Sequence 25, Appl
35	89.5	3.5	840	3	US-08-432-697-25	Sequence 25, Appl
36	89.5	3.5	840	3	US-08-466-248-25	Sequence 25, Appl
37	89.5	3.5	1013	4	US-08-860-886-2	Sequence 2, Appli
38	89.5	3.5	1250	1	US-08-441-139-9	Sequence 9, Appli
39	89	3.5	1014	3	US-09-078-347A-3	Sequence 3, Appli
40	89	3.5	1708	1	US-08-493-092-2	Sequence 2, Appli
41	89	3.5	1708	1	US-08-508-836A-2	Sequence 2, Appli
42	89	3.5	1786	3	US-08-973-462-8	Sequence 8, Appli
43	89	3.5	3056	1	US-08-508-836A-8	Sequence 8, Appli
44	89	3.5	3056	2	US-08-629-001A-3	Sequence 3, Appli
45	89	3.5	3056	2	US-08-874-266-2	Sequence 2, Appli

ALIGNMENTS

```

RESULT 1
US-08-928-862-3
; Sequence 3, Application US/08928862
; Patent No. 6309877
; GENERAL INFORMATION:
; APPLICANT: Chau, Raymond M. W.
; TITLE OF INVENTION: Isolation and Use of Motoneuronotrophic Factors
; FILE REFERENCE: 12592-2
; CURRENT APPLICATION NUMBER: US/08/928,862
; CURRENT FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-928-862-3

```

Query Match 100.0%; Score 2527; DB 4; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.8e-262;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 RAYYSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWWDIY 60
        |
Db      1 RAYYSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWWDIY 60

Qy     61 FNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEEL 120
        |
Db     61 FNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEEL 120

Qy    121 KLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDC 180
        |
Db    121 KLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDC 180

Qy    181 DPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDDSQICTVG 240
        |
Db    181 DPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDDSQICTVG 240

Qy    241 RLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTSVI 300
        |
Db    241 RLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTSVI 300

Qy    301 KIAVEPVNPSELPMKLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRMY 360
        |
Db    301 KIAVEPVNPSELPMKLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRMY 360

Qy    361 SEIDIKVADPVVTFCEVTVETSSLKCFATPNKKNKITMIAEPLKGLAEDIENEVQIT 420
        |
Db    361 SEIDIKVADPVVTFCEVTVETSSLKCFATPNKKNKITMIAEPLKGLAEDIENEVQIT 420

Qy    421 WNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQ 480
        |
Db    421 WNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQ 480

Qy     481 G 481
        |
Db     481 G 481
```

RESULT 2

US-09-293-549-2

; Sequence 2, Application US/09293549
; Patent No. 6440409
; GENERAL INFORMATION:
; APPLICANT: G. Todd Milne
; APPLICANT: Gerald Fink
; TITLE OF INVENTION: A METHOD FOR EFFICIENT AND HIGHLY
; TITLE OF INVENTION: SELECTIVE CONTROL OF MICROORGANISMS
; FILE REFERENCE: 50078/008002
; CURRENT APPLICATION NUMBER: US/09/293,549
; CURRENT FILING DATE: 1999-04-16
; EARLIER APPLICATION NUMBER: 60/082,089
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-293-549-2

Query Match 32.0%; Score 807.5; DB 4; Length 842;
Best Local Similarity 35.7%; Pred. No. 3e-77;
Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;

```
Qy      14 PLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIFYFNPKTRKFTKK-- 71
      | | | | | : ||: || || || :: : |||| :|||:|:| |
Db      204 PARGTVAFGSLHGWAFTRQFATRYAKKFG-VDKAKMMDRLWGDSSFFNPKTKKWTNKDT 262

Qy      72 -APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL 130
      | :|:| |||:|::: :: :| |:| | | :| | :|:|
Db      263 DAEGKPLERAFNMFILDPIFRLFTAIMNFKKDEIPVLEKLEIVLKGDEKDLEGKALLKV 322

Qy     131 VCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDLGEAMSDCDPDGFLMCHT 190
      | :|| :|| | | :|| :| | | | | |: :||| || :
Db     323 VMRKFLPAADALLEMIVLHLPSPVTAQAYRAEQLYEGPADDANCIAIKNCDPKADLMLYV 382

Qy     191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQICTVGRWLISVARYH 250
      :|| | | :|: |||| :||: || |:: | || :| | : |: :|:
Db     383 SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPGKKDDLFIAIKAIQRVVLMGRFV 442

Qy     251 IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
      :: |||| : : |:|| ::|| |:| :| | : :||: : |: :||| | :
Db     443 EPIDDCPAGNIIGLVGIDQFLLKTGTLTT---SETAHNMKVMKFSVSPVQVAVEVKNAN 499

Qy     311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADP 370
      :|||::||:|:| | | : | : ||||:|: ||||:|: : || : : :|:| :
Db     500 DLPKLVEGLKRLSKSDPCVLTYMSESGEHIVAGTGELHLEICLQDLEHDHAGVPLKISPP 559

Qy     371 VVTFCE TVVETSSSLKCF AETPNKKNKITMIAE PLEKGLAEDIENEV VQITWNRKKLGEFF 430
      || : ||| || :||| |:| : |||:|: :: ||| :: : |
Db     560 VVAYRETVESESSQTALSKSPNKHNRILYKAEPIDEEVSLAIENGINPRDDFKARARIM 619

Qy     431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
      | ||: || || |||| ||:::| | ||: | :||:|
Db     620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSVV 664
```

RESULT 3

US-09-293-549-4

; Sequence 4, Application US/09293549
; Patent No. 6440409
; GENERAL INFORMATION:
; APPLICANT: G. Todd Milne
; APPLICANT: Gerald Fink
; TITLE OF INVENTION: A METHOD FOR EFFICIENT AND HIGHLY
; TITLE OF INVENTION: SELECTIVE CONTROL OF MICROORGANISMS
; FILE REFERENCE: 50078/008002
; CURRENT APPLICATION NUMBER: US/09/293,549
; CURRENT FILING DATE: 1999-04-16
; EARLIER APPLICATION NUMBER: 60/082,089
; EARLIER FILING DATE: 1998-04-17

```
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-293-549-4
```

```
Query Match          32.0%; Score 807.5; DB 4; Length 842;
Best Local Similarity 35.7%; Pred. No. 3e-77;
Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;
```

```
Qy      14 PLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDIIYFNPKTRKFTKK-- 71
      | | | | | : | | | | | : : | | | | : | | | | | : | |
Db      204 PARGTVAFGSGLHGWAFITIRQFATRYAKKFG-VDKAKMMDRLWGDSFFNPKTKKWTNKDT 262

Qy      72 -APTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKHEELKLNIRPLLRL 130
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      263 DAEGKPLERAFNMFILDPIFRLFTAIMNFKKDEIPVLLKLEIVLKGDEKDLEGKALLKV 322

Qy      131 VCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHT 190
      | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db      323 VMRKFLPAADALLEMIVLHLPSVTAQAYRAEQLYEGPADDANCIAIKNCDPKADLMLYV 382

Qy      191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYLTLEDEEDSQICTVGRWLISVARYH 250
      : | | | | | | | | | | | | | | | | | | | | | | | | |
Db      383 SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPGKKDDLFIKAIQRVVLMMGRFV 442

Qy      251 IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
      : : | | | | : : | | | | : : | | | | : : | | | | : : | |
Db      443 EPIDDCPAGNIIGLVGIDQFLKLTGTLTT---SETAHNMKVMKFSVSPVVQVAVEVKNAN 499

Qy      311 ELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLD CVMHDLRKMYS EIDIKVADP 370
      : | | | : : | | | | | | | | | | | | | | | | | | | | |
Db      500 DLPKLVEGLKRLSKSDPCVLTYMSESGEHIVAGTGELHLEICLQDLEHDHAGVPLKISPP 559

Qy      371 VVTFCE TVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIENE VVQITWNRKKLGEFF 430
      | | : | | | | | : : | | | | | | : : | | | | : : | |
Db      560 VVAYRETVESESSQTALSKSPNKHNRILYKAEPIDEEVSLAIENGIINPRDDFKARARIM 619

Qy      431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
      | | | : | | | | | | | | | | | | | | | | | | | | |
Db      620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSVV 664
```

RESULT 4

US-09-183-861-85

; Sequence 85, Application US/09183861

; Patent No. 6365165

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND

; NUMBER OF SEQUENCES: 87

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,861
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/022,765
; FILING DATE: 12-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.420C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-183-861-85

```

```

Query Match          21.9%;  Score 554.5;  DB 4;  Length 361;
Best Local Similarity 34.6%;  Pred. No. 1e-50;
Matches 122;  Conservative 82;  Mismatches 136;  Indels 13;  Gaps 6;

```

```

Qy      66 RKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVTSLPRTLDELGIHLTKEELKL 122
      :|: |:  :  :  |:| :| |:|:|:|  |:  :  :  |  |  :| |  :
Db      1  KKWKQETNADGERVRR AFCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQ 60

Qy     123 NIRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS D----LGEAMS 178
      ||: |  |  :  | | |:| |  :  |  |:|  :  :  :|  :
Db     61 VPXKLLKTVMXFLPAAETLLQMI VAHLPS PKKAQAYRAEMLYSGEASPEDKYFMG--IK 118

Qy     179 DCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTL EDEED-SQIC 237
      :| |  ||| : :| |  :|  |||: || : :| |:|:| |  :|  :
Db     119 NCDPAAPLMLYISKMVP TADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDK 178

Qy     238 TVGRLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGN EEAQIFRPLKFNTT 297
      | |  : : | |  :| | | : |||: |||:| | |  :  | :| : :
Db     179 PVQRSVLMMGRYQEAVEDMPCGNVVGVLGVVDKYIVKSATITD--DGESPHPLRDMKYSVS 236

Qy     298 SVIKIAVEPVNPSEL PKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLR 357
      |:|:| | | |:| |:|:| |:|:| | | :  :| | | |:| : | | :
Db     237 PVVRVAVEAKNPSDLPKLVEGLKRLAKSDPLVVC SIEESGEHIVAGAGELHLEICLKDLQ 296

```

Qy 358 KMY-SEIDIKVADPVVTFCEVTVVETSSSLKCFEAETPNKKNKITMIAEPLKGLA 409
 : : : :|:::| | | : | | : | : : | | : : | | : | |
 Db 297 EDFMNGAPLKISEPVVVSFRETVDVSSQQCLSKSANKHNRLFCRGAPLTEXLA 349

RESULT 5

US-09-022-765-85

; Sequence 85, Application US/09022765

; Patent No. 6375955

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/022,765

; FILING DATE: 12-FEB-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.420C3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 85:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 361 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-09-022-765-85

Query Match 21.9%; Score 554.5; DB 4; Length 361;

Best Local Similarity 34.6%; Pred. No. 1e-50;

Matches 122; Conservative 82; Mismatches 136; Indels 13; Gaps 6;

Qy 66 RKFTKKAPTSSSQ---RSEVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKL 122

:|: |: : : |:| :| |:|:|:| |: : : :| |:| |:| :

Db 1 KKWIKQETNADGERVRR AFCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQ 60

Qy 123 NIRPLLRVLCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSD----LGEAMS 178
 Db 61 VPXKLLKTVMXFLPAAETLLQMI VAHLPSPKKAQAYRAEMLYSGEASPEDKYFMG~-IK 118

Qy 179 DCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED-SQIC 237
 Db 119 NCDPAAPLMYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNVYVGKKQDLYEDK 178

Qy 238 TVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
 Db 179 PVQRSVLMMGRYQEAVEDMPCGNVVGVLGVGDKYIVKSATITD--DGESPHPLRDMKYSVS 236

Qy 298 SVIKIAVEPVNPSELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLR 357
 Db 237 PVVRVAVEAKNPSDLPLKLV EGLKRLAKSDPLVVC SIEESGEHIVAGAGELHLEICLKDLQ 296

Qy 358 KMY-SEIDIKVADPVVTF CETVVETSS LKCF AETPNKKNKITMIAEPLEKGLA 409
 Db 297 EDFMNGAPLKISEPVVSFRET VTDVSSQCLSKSANKHNRLFCRGAPLTEXLA 349

RESULT 6

US-09-551-974A-85

; Sequence 85, Application US/09551974A

; Patent No. 6500437

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

; FILE REFERENCE: 210121.420C5

; CURRENT APPLICATION NUMBER: US/09/551,974A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 101

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 85

; LENGTH: 361

; TYPE: PRT

; ORGANISM: Leishmania major

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(361)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-551-974A-85

Query Match 21.9%; Score 554.5; DB 4; Length 361;

Best Local Similarity 34.6%; Pred. No. 1e-50;

Matches 122; Conservative 82; Mismatches 136; Indels 13; Gaps 6;

Qy 66 RKFTTKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKL 122
 Db 1 KKWIKQETNADGERVRR AFCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQ 60

Qy 123 NIRPLLRVLCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSD----LGEAMS 178


```

      ||: | | : | | | : || | : | : | :
Db      61 VPXKLLKTVMMXFLPAAETLLQMIVAHLPSPKKAQAYRAEMLYSGEASPEDKYFMG--IK 118

Qy      179 DCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED-SQIC 237
      :||| ||| : :|| | | :| |||: || : :|| |:::| || :|| :
Db      119 NCDPAAPLMPLYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDK 178

Qy      238 TVGRLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
      | | : : || | :| || | : |||: |||:||||: | | :|:::
Db      179 PVQRSVLMMGRYQEAVEDMPCGNVVGLVGVDKYIVKSATITD--DGSEPHPLRDMKYSVS 236

Qy      298 SVIKIAVEPVNPSELPPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLR 357
      |:::| |||:||||:|:::| || | : :| |||:|: | |||:|: : ||:
Db      237 PVVRVAVEAKNPSDLPKLVEGLKRLAKSDPLVVCSIEESGEHIVAGAGELHLEICLKDLQ 296

Qy      358 KMY-SEIDIKVADPVVTFCETVVETSSLKCFETPNKKNKITMIAEPLEKGLA 409
      : : : :|:::| ||| : || :| ::: || |:: || : ||
Db      297 EDFMNGAPLKISEPVVSFRETVDVSSQQCLSKSANKHNRLFCRGAPLTEXLA 349

```

RESULT 7

US-09-252-991A-25011

; Sequence 25011, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25011

; LENGTH: 900

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (857)

; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-25011

Query Match 7.1%; Score 179; DB 4; Length 900;

Best Local Similarity 22.2%; Pred. No. 1e-09;

Matches 92; Conservative 55; Mismatches 143; Indels 124; Gaps 18;

```

Qy      109 DEL-----GIHLTKHEELKLNIRPLLRLVCKKFFGEFTG-----FVDMCVQHIPS 152
      ||| | | :| |||: | | : : | | :| : :||:
Db      386 DELMNKYLEGEELSIEEIKAGVRQ--RTLANQIVPAVLGSSFKNGVPLVLDAVIDYLP 443

Qy      153 PKVGAKPKIEHTYTGGVSDSLGEAMSD--CDPDGPLMCHTTKMFSTHDGVQFHPFGRVLS 210

```

```

Db      444 PS-----EIPAIRGTDPDDEEKHDERHADDDPEFSSALAFKI-ATDPFVGTLT'FARVYS 495
QY      211 GTIHAGQPV--KVLGENYTLDEEDDSQICTVGRLWISVARYHIEVNRVPAGNWWLIEG-- 266
      | : : | | | : | | : | | : | | : |
Db      496 GVLTSGLDAVLNSVKGKKE-----RVGRMVQMCHANQRDEIKEVRAGDIAALIGMK 544
QY      267 -----VDQPIV-KTATITEPRGNEEAQIFRPLKFNTTTSVIKIAVEPVNPSELPKM 315
      : | : | : : | | : | | | : : | |
Db      545 DVTTGDTLCAIDKPIILERMDFPDP-----VISVAVEPKTKADQEKM 586
QY      316 LDGLRKVNKSYPSLTKV-EESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPVVTF 374
      | | : : | | | : : : | | | : | : : : : | | :
Db      587 GIALSKLAQEDPSFRVKTDEETAQTIISGMGELHLDIIVDRMRREFG-VEANIGKPQVAY 645
QY      375 CETVVETSSLKCFEAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFFQTKY 434
      | | : | | | : | | : | | : | : | : |
Db      646 RETIRNT----C-----EIEGKFVRQSGGRGQFGH----- 671
QY      435 DWDLLAARSIW-AFGPDATGP-----NILVDDTLPSEVDKALLGSKVDSIVQG 481
      | | | | | : | : | | | : : | : |
Db      672 -----CWIRFAPADEGQEGLEFHNEVVGGVIPREFIPAIQKGIEDQMONG 716

```

RESULT 8

US-09-252-991A-21345

; Sequence 21345, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21345

; LENGTH: 716

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21345

Query Match 7.0%; Score 177; DB 4; Length 716;

Best Local Similarity 21.6%; Pred. No. 1.1e-09;

Matches 99; Conservative 75; Mismatches 172; Indels 112; Gaps 21;

```

QY      59 IYFNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKE 118
      | | : | : : : : | : : : | | | :
Db      201 IYWNDDDKGMTYREEEIPAE---LKDLAEWRSSMVEAAAEANEELMNKYLEEG-ELSEA 256
QY      119 ELKLNIRPLLRLVCK-----KFFGEFTGFV-DMCVQHIPSQVGAQPKIEHTYTG 169
      | : | : | | : | : | : : | : |

```

Db 257 EIKEGLR-LRTLACEIVPAVCGSSSEFKNKGVLVLDAVIDYLPAPT--EIPAIAK-----GV 308

QY 170 DSDLGEAMSD---CDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPV--KVLGE 224
| | : | | : | | : : | | | | : : | | | :

Db 309 SPD-DETVEDERHADDNEPFSSLAFKI-ATDPFVGTLTFARVYSGVLSSGDSVLNSVKGK 366

QY 225 NYTLEDEEDSQICTVGRWLISVARYHIEVNRVPAGNWVLIEG-----VDQPIV 272
| | : | | : | | : : | | : : | | :

Db 367 KE-----RVGRMVQMCHANQREEIKEVRAGDIAALIGMKDVTGTGDTLCSIEKPII 415

QY 273 -KTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSELPMKMLDGLRKVNKSYPSLTT 331
: | | | : | | | : | | : : | | | : : | |

Db 416 LERMDFPEP-----VISVAVEPKTKADQEKMGIALGKLAQEDPSFRV 457

QY 332 KV-EESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADPVVTFCTVETSSLKCFEAT 390
| | | | : : | | | | : : : : : : : : | | : | | : : :

Db 458 KTDEESGQTIISGMGELHLDIIVDRMKREFG-VEANIGKPQVAYRETITKDN----- 508

QY 391 PNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPD 450
: | | : | : : | : | : | : | : |

Db 509 -----VEIEGKFVRQSGGRGQFG-----HCWIRFSAADVDEKGN 543

QY 451 ATG---PNILVDDTLPSE----VDKALLGSVKDSIVQG 481
| | : | : | : | : : | : : | |

Db 544 TEGLVFENEVVGGVVPKEYIPAIQKGIEEQMKNGVVAG 581

RESULT 9

US-09-134-001C-5356

; Sequence 5356, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5356

; LENGTH: 705

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5356

Query Match 6.9%; Score 175; DB 4; Length 705;

Best Local Similarity 23.9%; Pred. No. 1.8e-09;

Matches 92; Conservative 70; Mismatches 151; Indels 72; Gaps 17;

QY 71 KAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL 130
| : : | : : : : | : | : : | : |

Db 216 KERAEEARAQLIEAVAENNDLMEKYLGDDEEIS---VDELKDAIRQATTDVEFYFVL-- 269

Db 199 TEDEIHQVMRKGVIENTKINPVL---CGTAFKNKGVQQLNVIVKWLPSP-----LDRG 248

Qy 165 YTGGVSDSLGEAMSDCDP--DGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVL 222
| : : : | : | | | | : | | | : | : | : | : | : |

Db 249 NIRGINLKTDQEIS-LEPRRDGFLAALAFKIM-TDPYVGRITFIRIYSGTLKKGSAIL-- 304

Qy 223 GENYTLEDEEDSQICTVGRWLISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRG 282
| | : : | : | | : : | : | : : |

Db 305 --NSTKDKKE-----RISRLLEMHANERTDRDEFTVGDIGACVGLKFSVTGDTLCDD--- 354

Qy 283 NEEAQIFRPLKFNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTTKV-EESGEHVI 341
| : | : | : | : | : | : | : | : | : | : | : | : | : |

Db 355 NQEIVLER-IEF-PDPVIDMAIEPKSKGDREKLAQALSSLSEEDPTFRVSTNEETGQTII 412

Qy 342 LGTGELYLDCVMHDLRKMYSIDIKVADPVVTFCEVTVETSSLKCFEAETPNKKNK----- 396
| | | : | : | : | : | : | : | : | : | : |

Db 413 SGMGELHLD-ILRDRMIREFKVEANVGKPVSYKETITVSGN---SETKYVKQSGGRGQ 467

Qy 397 ---ITMIAEPLKGLAEDIENEVV 417
: : | | | : : : |

Db 468 YAHVCLEIEPNEPGKGNEVVSKIV 491

RESULT 11

US-09-328-352-4749

; Sequence 4749, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4749

; LENGTH: 277

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4749

Query Match 4.5%; Score 113.5; DB 4; Length 277;

Best Local Similarity 21.3%; Pred. No. 0.0015;

Matches 35; Conservative 33; Mismatches 57; Indels 39; Gaps 4;

Qy 322 VNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADPVVTFCEVTVET 381
: | : | | | : | | | : | : : : : | : | : | : |

Db 4 LKKTFFHSGVHTDEESGQTIIAGMGELHLDIIVDRMKREFG-VEANIGKPMVAYRETIKKT 62

Qy 382 SSLKCFEAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFFQTKYDWDLLAA 441
: | : | : | : | : | : |

Db 63 -----VEQEGKFVRQTGGKGKFGHVYVRLEPLDVEAA 94

Qy 442 RSIWAFGPDATGPNILVDDTLPSE----VDKALLGSVKDSIVQG 481
: | : | : | | | : | : : |

Db 95 GKEYEFAEEVVG-----GVVPKEFFGAVDKGIQERMKNGVLAG 132

US-09-252-991A-19697

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

; FILE REFERENCE: 107196.136

```

; CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

: PRIOR FILING DATE: 1998-02-18

: PRIOR APPLICATION NUMBER: US 60/094,190

: PRIOR FILING DATE: 1998-07-27

: NUMBER OF SEQ ID NOS: 33142

: SEO ID NO 19697

```

:      SEQ ID NO 1569
:      LENGTH: 1067

```

```

; LENGTH: 1
; TYPE: PRT

```

```

: ORGANISM: Pseudomonas aeruginosa

```

US-09-252-991A-19697

Query Match 4.4%; Score 111; DB 4; Length 1067;

Best Local Similarity 22.5%; Pred. No. 0.028;

Matches 96; Conservative 65; Mismatches 132; Indels 134; Gaps 28;

```

Qy      68 FTKKAPTSSSSQRSFVEFILEPLYKILAQ--VVGVDVDTSLPRTLDELGIHLTKHEELKLNIR 125
      ||:  |:  ::| |:| |:| | :||  | | ::::  ||  |:
Db      630 FTRLRATA----DWLEQIAKNLTKLLAQHPLPGD----LGR LIEQV-----PELAREIK 675

```

Qy 126 ---PLLRLVCKKFFGEFTGFVDMCVQHIPS PKV--GAKPKIEHTYTGGVD-----SD 172
: |:: |:| || : | : | | || |:: :|
Db 676 TQQQFMFTACEE-IGDFRAGEDMEGRERPRHRFVGGVVP--EHIREMGIELKKGF SKLTD 732

```

QY      173 LGEAMSD-----CDPDGPLMCHTTKMFSTHDGVQFHP-FGRVLSGTIHAGQPVKVLGENY 226
      |  ::|          | :|          :|  ::| || :|:          |  |  :
Db      733 LFTRLTDILKEAMDGEG-----AGGIASHQAEWYPLFGSLA----RAQGNWELWTA 782

```

```

Qy      227 TLEDEEDSQICTVGRWLWISVAR----YHIEVNRVP-----AGNWVLI EGVDQPIVKT 274
          | | | : | | : | | : : | | | | | | | | | | | : | | : | :
Db      783 TCEDPDQDSP--PMAR-WLT LAESGSFYDIEANASPILAAETLRRNLWNVAYGV---LVTS 836

```

```

Qy      275 ATITE-----PRGNEEAQIFRPLKFNTTSVIKI---AVEPVNPS----- 310
      ||:|          ||      | : |          |::      :| | :
Db      837 ATLTALGTFDRYMRAGLPRNAVTAVVSPFHHAAEAGVLRVPDLKADPRNAAEHTAAIIR 896

```

Qy 311 ELPKMLDGLR-----KVNKSYPSL-----TTKVEESGEHVI-LGTGE 346
|||:::| | ::::| :| | :| :||
Db 897 ELPELVKGARGSLVLFSSRKQMQEVFDGLDRDWRKRVFIQGNLSKQETLNKHKSRVDDGE 956

```

Qy      347  LYLDCVMHDLRKMYSEIDIKVADPVVTFCETVVETSSLKCFAETPNKNKITMIAEPLEK 406
          |  | : |           : | : | | |           | : |           : : | : |
Db      957  ---DSVLFGLASFAEGVDLPGA-----YCEHVV-----IAKIP-----FAVPDDPVEA 996

```

Qy 407 GLAEDIE 413
||| ||
Db 997 ALAEWIE 1003

RESULT 13

US-09-107-532A-6220

; Sequence 6220, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND

THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 6220:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1224 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (B) LOCATION 1...1224

; SEQUENCE DESCRIPTION: SEQ ID NO: 6220:

US-09-107-532A-6220

Query Match 4.2%; Score 105; DB 4; Length 1224;
 Best Local Similarity 19.4%; Pred. No. 0.15;
 Matches 100; Conservative 63; Mismatches 171; Indels 182; Gaps 25;

```

Qy      15 LLGNVCFSSSQYSICFTLGSEFAKIY-----ADTFGDINYQEFAKRLWGDIIYFNPKT 65
      |||  | : | : | | : | : | : | : | : | : |
Db      338 LLGKRVRHYSGRSVI--VVGPFLLKMYQCGLPKEMAIELFKPFVMRELVRQELASNIKNAK- 394

Qy      66 RKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPR--TLDELGIHLTKEEL--- 120
      ||  ::  : :: : | | : :  | |  ||  |||  : |
Db      395 RKIERQE-----DEVVDVLEDDVIKEHPVLLNRAPTLHRLGIQAFEPVLVQG 440

Qy      121 -KLNIRPLRLVCKKFFGEFTGFVDMCVQHIPP-----SPKVGAKP 159
      : : | ||| : : : | | | : | : || | ||
Db      441 RAIRLHP---LVCEAYNADFDG--DQMAVHVPLNEEAQAEARMLMLAAQNILNPKDG-KP 494

Qy      160 KIEHTYTGGVDSDLGEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVL---SGTIHA 215
      :  |  |  ||  : : | : : | : :  : : | : : |
Db      495 VV---TPSQDMVLGNYYLTMEEEG-----REGEGMVFRDQNEAVIAWRNGYVHL 540

Qy      216 GQPVKV---LGEN-YTLEDEEDSDQICTVGRLWISVA-----R 248
      : |  ||| : | : | : | ||| : :
Db      541 HSRIGVNPNSLGEKPFTEWQKERTMITTVGKII FNEIMPPEFFPYLNEPTDFNLTVQTPDK 600

Qy      249 YHIEVNRVPAGNWVLIEGVDQPI-VKTATITEP-----RGNEEAQIFRPLKFNTTSVIKI 302
      | : |  | | | : | : |  ||  | : : | : |  ||
Db      601 YFVEA-----GTDIPAHIKEQELVLPFKKKNLGNIIAEVFKRFKVTETS----- 644

Qy      303 AVEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSE 362
      |||| :: : : :
Db      645 -----KMLDRMKDLGYKHST-----HAG 662

Qy      363 IDIKVAD-PVVTFCETVVETSSSLKCFEAETPNKKNKITMIAEPLKGLAEDIEN-EVVQIT 420
      | : : || | : : : | | : : | : : || | | |
Db      663 ITVGIADISVLNEKQEIIIE-----NAHKQVETITKQFRRLITDDERYERVIAV 711

Qy      421 WNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNI 456
      || | : | | | | : |
Db      712 WNAAK--DSIQQKLMEGLEAKNPIFMMSDSGARGNI 745

```

RESULT 14

US-09-328-352-8132

; Sequence 8132, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 8132

; LENGTH: 1380

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii
US-09-328-352-8132

Query Match 4.1%; Score 103; DB 4; Length 1380;
Best Local Similarity 19.5%; Pred. No. 0.31;
Matches 123; Conservative 99; Mismatches 192; Indels 218; Gaps 32;

```
QY      25 QYSICFT----LGSFAKIYADTF-----GDINYQEFARLWGLDI 59
      |::| ||      |||      |||      |:  :| |: |: ::
Db      351 QFNILFTNDIDRGSFV---ADTLRADLTRDREEALVEIYKVMRPGEPPTKEAAENLFNNL 407

QY      60 YFNPK-----TRKFTKKAPTSSSQRS-FVEFIL--EPLYKILAQVV-----G 98
      :|: :      |: :      : |:| || || |: :| :|      |
Db      408 FFSSERYDLSPVGRMKFNRRRLGRPYEVGTDQKSREVEGILSHEDIIDVLRTLVEIRNGKG 467

QY      99 DVDTSLPRTLDELG-----IHLTKEELK-----LNI 124
      :||      :| ||      |||      :|
Db      468 EVD-----DIDHLGNRRVRSVGEMTENQFRVGLVRVERAVKERLSQAETDNLSPODLINA 522

QY     125 RPLRLVCKKFFG--EFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGV----- 169
      :|: : |:||| : : |:| |: | : : | : : ||:
Db     523 KPVAAAI-KEFFGSSQLSQFMD---QNNPLSEITHKRRVSALGPGGLTRERAGFEVRDVH 578

QY     170 DSDLGEAMSDCDPDGFL--MCHTTKMFSTHDGVQF--HPFGRVLSGTI----- 213
      : |      |:|| : : : : : | | : :|: | :
Db     579 QTHYGRVCPIETPEGPNIGLINSLSVYAKANDFGFLETPIRKVVDGRVTDDVEYLSAIEE 638

QY     214 -----HAGQPVKVLGENYTLDEEDSQICTVGRLWISVA 247
      | |: |: : | | | | : :| |
Db     639 VGTVIAQADS AVDKDGNLTEEFVSVRHQGEFVRMPPEKVTHMDVSAQQQVVSVAASLIPFL 698

QY     248 RYHIEVNRVPAGNWVLIEGV-----DQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKI 302
      | : || |: : : | |:|:| | || : | : | |
Db     699 E-HDDANRALMGSNMQRQAVPTLRADKPLVGTGM-----EANVAR-----DSGVCVI 744

QY     303 A-----VEPVNPSEL-----PKMLDG-----LRKVNSYPSLTTKVEES-----GE 338
      | : | |: | : : |: | | | : | : : : | :
Db     745 ANRGGVIEYVDASRIVIRVNEDEM VAGEAGVDIYNLIK YTRS--NQNTCINQNVIVNLGD 802

QY     339 HVILGTGELYLDCVMHDLRKMYS EIDIKVADPVVTF CETVVETSSLKCF AETPNKKNKIT 398
      | |: : | |: : : :|| :|: | | | :| : : :|
Db     803 KV--ARGDILADGPSTDMGELALGQNM RVA--FMTWNGYNYEDSIL--LSERVLQEDRLT 856

QY     399 MIAEPLEKGLAEDIENEVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILV 458
      | :| | : :|| : :|| :| | | : | :| |
Db     857 SIHIQELSCVARDTKLGAEEITADIPNVGEAALSKLD-----ESGIVYIGA EVTAGDILV 911

QY     459 DDTLP-----SEVDKALLG----SVKDS 477
      | : : |: | |||
Db     912 GKVTPKGETQLTPEEKLLRAIFGEKAADV KDS 943
```

RESULT 15

US-09-252-991A-32924
; Sequence 32924, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 11:21:12 ; Search time 26.2023 Seconds
(without alignments)
1765.382 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSVKDSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1764.5	69.8	974	2	T29007	translation elonga
2	1528.5	60.5	987	2	H86197	hypothetical prote
3	1335	52.8	983	2	T39902	translation Elonga
4	870	34.4	858	1	EFHU2	translation elonga
5	868	34.3	858	1	EFRT2	translation elonga
6	863	34.2	858	2	A25440	translation elonga
7	845.5	33.5	844	2	S05988	translation elonga
8	837.5	33.1	852	2	T21362	hypothetical prote
9	827.5	32.7	852	2	A40411	translation elonga
10	816.5	32.3	843	2	T14579	translation elonga
11	807.5	32.0	842	2	A41778	translation elonga
12	801.5	31.7	842	2	T41697	translation elonga
13	793.5	31.4	846	2	A96602	elongation factor

14	787	31.1	848	2	G90128	elongation factor
15	779	30.8	845	2	S32819	translation elonga
16	710.5	28.1	830	2	A34347	translation elonga
17	668.5	26.5	1008	2	S38003	translation elonga
18	646	25.6	361	2	S07567	translation elonga
19	458.5	18.1	1000	2	T41396	probable translati
20	404.5	16.0	736	2	G72621	probable translati
21	396.5	15.7	735	2	G71203	probable translati
22	390.5	15.5	728	2	E69486	translation elonga
23	390.5	15.5	1110	2	S60964	probable membrane
24	385.5	15.3	732	1	S23864	translation elonga
25	385.5	15.3	732	2	A75219	translation elonga
26	366.5	14.5	736	1	S31809	translation elonga
27	347.5	13.8	179	2	T03215	translation elonga
28	346	13.7	725	2	T44992	translation elonga
29	335	13.3	737	1	S14408	translation elonga
30	331	13.1	726	1	G64430	translation elonga
31	326.5	12.9	732	1	S36089	translation elonga
32	326.5	12.9	733	2	E69007	translation elonga
33	322.5	12.8	734	1	S25166	translation elonga
34	318.5	12.6	727	1	S01289	translation elonga
35	309	12.2	682	2	B84415	translation elonga
36	305.5	12.1	730	2	T43943	translation elonga
37	303	12.0	728	1	S07558	translation elonga
38	295	11.7	730	2	T44246	translation elonga
39	285	11.3	730	2	T44066	translation elonga
40	269	10.6	906	2	T23556	hypothetical prote
41	225.5	8.9	698	2	E75536	translation elonga
42	219	8.7	695	2	AE1406	translation elonga
43	216	8.5	695	2	AE1782	translation elonga
44	207.5	8.2	827	2	G90092	hypothetical prote
45	206	8.2	695	2	S76751	translation elonga

ALIGNMENTS

RESULT 1

T29007

translation elongation factor eEF-2 homolog eft-1 [similarity] - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Nov-2000

C;Accession: T29007

R;Favella, A.

submitted to the EMBL Data Library, March 1996

A;Description: The sequence of *C. elegans* cosmid ZK328.

A;Reference number: Z20552

A;Accession: T29007

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-974 <FAV>

A;Cross-references: EMBL:U50193; PIDN:AAA91248.1; CESP:eft-1

A;Experimental source: strain Bristol N2

C;Genetics:

A;Gene: CESP:eft-1

A;Introns: 210/1; 335/1; 686/2; 876/3

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

F;133-264/Domain: translation elongation factor Tu homology <ETU>

Query Match 69.8%; Score 1764.5; DB 2; Length 974;
Best Local Similarity 69.1%; Pred. No. 1.1e-122;
Matches 331; Conservative 66; Mismatches 81; Indels 1; Gaps 1;

```
QY      4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGD-INYQEFARLWGDYFN 62
      :: :: :|:|:| | | | | | :|:|:|:| | | :| | | | | :|:|:|:|:|:|
Db      295 FAEDVPVLSPLNGNVIFSSGRYNVCFSLLSFSNIYAKQHGDSEFNSKEFARRLWGDYFE 354

QY      63 PKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKL 122
      | | | | | | :|:|:| : |:|:|:|:|:|:|:|:|:|:|:| | | : | | | | | | :|
Db      355 KKTRKFVKKSPSHDAPRTFVQFILEPMYKIFSQVVGVDVDTCLPDVMAELGIRLSKEEQKM 414

QY      123 NIRPLLRVLCKKFFGEFTGFVDMCVQHIPSQVKGAKPKIEHTYTGGVSDLGEMSDCDP 182
      |:|:|:| |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      415 NVRPLIALICKRFFGDFSAFVDLVVQNIKSPLNNAKTKEQTYLGPADSQLAQEMQKCNA 474

QY      183 DGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRL 242
      :|:|:| | | | | : | | | | |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      475 EGPLMVHTTKNYPVDDATQFHVFGVRMSGTLEANTDVRVLGENYSIQDEEDCRRMTVGRL 534

QY      243 WISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKI 302
      :: | | | | |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      535 FVRVASYQIEVSRVPAGCWVLIEGIDQPIVKTATIAELGYEEDVYIFRPLKFNTTRSCVKL 594

QY      303 AVEPVNPSELPMKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDLCVMHDLRKMYSE 362
      | | | |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      595 AVEPINPSELPMKMLDGLRKVNKSYPLLTTRVEESGEHVLLGTGEFYMDLCVMHDMRKVFSE 654

QY      363 IDIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLEKGLAEDIENEVVOITWN 422
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      655 IDIKVADPVVTFNETVIETSTLKCAETPNKKNKITMMAEPLEKQLDEDIENEVVOIGWN 714

QY      423 RKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSVDKALLGSVKDSIVQG 481
      |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      715 RRRLGEFFQTKYNWDLLAARSIWAFGPDTTGPNILLDDTLPSVDKHLSTVRESLVQG 773
```

RESULT 2

H86197

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: H86197

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;
Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;
Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;
Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;
Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.;
Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;
Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-
Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu,

S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militsher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-987 <STO>
A;Cross-references: GB:AE005172; NID:g8844127; PIDN:AAF80219.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

Query Match 60.5%; Score 1528.5; DB 2; Length 987;
Best Local Similarity 61.2%; Pred. No. 3.8e-105;
Matches 289; Conservative 69; Mismatches 113; Indels 1; Gaps 1;

```

Qy      11  ILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFG-DINYQEFARLWGDIFYNPKTRKFT 69
          :: | ||||:| || ||||:| | :: :|| ||||:|:: || |
Db      312 LIDPAAGNVCFASGTAGWSFTLQSFARMYAKLHGVAMDVDKFAASRLWGDVYYHSDTRVFK 371

Qy      70  KKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLNIRPLLR 129
          : | :||:||||||| :||:| | || ||: | : |||||
Db      372 RSPPVGGGERAFVQFILEPLYKIYSQVIGEHKKSVEVTTLAELGVTLSNSAYKLNVRPLLR 431

Qy     130  LVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCH 189
          | | || :|| || |:||||: | |:::|| | | : | :|| |||| :
Db     432 LACSSVFGSASGFTDMLVKHIPS PREAAARKVDHSYTGTKDSP IYESMVECDPSGPLMVN 491

Qy     190  TTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEEDSQICTVGRLWISVARY 249
          ||:: | | |||| || : || |:|||| | : |||| | | :|| |||
Db     492 VTKLYPKSDTSVFDVFGRVYSGRLQTGQSVRVLGEGYSPEDDEEDMTIKEVTKLWIYQARY 551

Qy     250  HIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNP 309
          | | : | |:||||| |:||||: :| : || |:|| | :| | ||:|
Db     552 RIPVSSAPPGSWVLIEGVDASIMKTATLCNASYDEDVYIFRALQFNTLPVVKTATEPLNP 611

Qy     310  SELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRMYSEIDIKVAD 369
          |||||:|:|||:| || | ||||| | ||||| | : | |||:|:|||
Db     612 SELPKMVEGLRKISKSYPLAITKVEESGEHTILGTGELYLDSIMKDLRELYSEVEVKVAD 671

Qy     370  PVVTFCEVTVETSSLKCFEAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEF 429
          |||:||||||:|:||||| ||||| ||||| || | ||||:|:|
Db     672 PVVSFCETVVESSMKCFAETPNKKNKITMIAEPLDRGLAEDIENGVSIDWNRKQLGDF 731

Qy     430  FQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPEVDKALLGSVKDSIVQG 481
          |:||||| ||||| |||||:||||:|: | : |||||
Db     732 FRTKYDWDLLAARSIWAFGPDQGNILLDDTLPEVDRNLMMAVKDSIVQG 783

```

RESULT 3

T39902
translation Elongation Factor 2 - fission yeast (*Schizosaccharomyces pombe*)
C;Species: *Schizosaccharomyces pombe*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Sep-2000
C;Accession: T39902
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z21889
A;Accession: T39902
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-983 <LYN>
A;Cross-references: EMBL:AL033534; PIDN:CAA22126.1; GSPDB:GN00067;
SPDB:SPBC215.12
A;Experimental source: strain 972h-; cosmid c215
C;Genetics:
A;Gene: SPDB:SPBC215.12
A;Map position: 2
A;Introns: 8/2; 128/2; 245/3
C;Superfamily: translation elongation factor 2; translation elongation factor Tu
homology

Query Match 52.8%; Score 1335; DB 2; Length 983;
Best Local Similarity 53.7%; Pred. No. 8.7e-91;
Matches 257; Conservative 74; Mismatches 146; Indels 2; Gaps 1;

Qy	5	STDENLILSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFKRLWGDYFNPK	64
		: : : : :	
Db	304	SKDLKYRVSPELGNVCFASCDLGYCFTLSSFAKLYIDRHGGIDVDLFSKRLWGDYFDSK	363
Qy	65	TRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKLN	124
		: : : : : : : :	
Db	364	TRKFAKQSLDGSQVRSFVHFILEPLYKLHTLTISDEAEKLLKHLSSFQIYLPKPDYLLDP	423
Qy	125	RPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDC--DP	182
		: : : : : : : : : : :	
Db	424	KPLLQLICASFFGFVGVNAVTRHIPSRENAARKASQSYIGPINSSIGKAILEMSREE	483
Qy	183	DGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQICTVGRL	242
		: : : : : : : : :	
Db	484	SAPLVMHVTCLYNTVDANNEYAFARVYSGQVKKGQKVKVLGENYSLEDEEDMVAHIAEI	543
Qy	243	WISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNNTTSVIKI	302
		: : : : : : : :	
Db	544	CVPCARYRLHVDGAVAGMLVLLGGVDNSISKATATIVSDNLKDDPYIFRPIAHMSSESVFKV	603
Qy	303	AVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSE	362
		: : :	
Db	604	AVEPHNPSELPKLLDGLRKTNKSYPSTTKVEESGEHTIFGTGEMYMDCLLYDLRTLYSE	663
Qy	363	IDIKVADPVVTFCEVTVETSSILKCFATPNKKNKITMIAEPLEKGLAEDIENEVQITWN	422
		: : : : : : : :	
Db	664	IEIRVSDPVARFCETAVDTSSIKCFSDTPNKKNRITMVVEPLEKGISNDIENGKVNINWP	723
Qy	423	RKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG	481
		: : : : :	
Db	724	QKRISSEFFQKNYDWDLLASRSIWAFGPDDRGTNILRDDTLSTDVDKNVLSVKEYIKQG	782

RESULT 4

EFHU2

translation elongation factor eEF-2 - human

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence_revision 03-Oct-1995 #text_change 19-Jan-2001

C;Accession: S18294; S00467; H33178; S06137

R;Rapp, G.; Klaudiny, J.; Hagendorff, G.; Luck, M.R.; Scheit, K.H.

Biol. Chem. Hoppe-Seyler 370, 1071-1075, 1989

A;Title: Complete sequence of the coding region of human elongation factor 2 (EF-2) by enzymatic amplification of cDNA from human ovarian granulosa cells.

A;Reference number: S06137; MUID:90121741; PMID:2610926

A;Accession: S18294

A;Molecule type: mRNA

A;Residues: 1-858 <RAP>

A;Cross-references: EMBL:X51466; NID:g31105; PIDN:CAA35829.1; PID:g31106

R;Rapp, G.; Mucha, J.; Einspanier, R.; Luck, M.; Scheit, K.H.

Biol. Chem. Hoppe-Seyler 369, 247-250, 1988

A;Title: Cloning and sequence analysis of a cDNA from human ovarian granulosa cells encoding the C-terminal part of human elongation factor 2.

A;Reference number: S00467; MUID:88293714; PMID:2840927

A;Accession: S00467

A;Molecule type: mRNA

A;Residues: 501-858 <RA2>

A;Cross-references: EMBL:M19997; NID:g181968; PIDN:AAA50388.1; PID:g181969

A;Note: the amino end of the mature protein was determined by protein sequencing

R;Ward, L.D.; Hong, J.; Whitehead, R.H.; Simpson, R.J.

Electrophoresis 11, 883-891, 1990

A;Title: Development of a database of amino acid sequences for human colon carcinoma proteins separated by two-dimensional polyacrylamide gel electrophoresis.

A;Reference number: A33178; MUID:91176935; PMID:2079031

A;Accession: H33178

A;Molecule type: protein

A;Residues: 2-14,'X',16,'X',18-19 <WAR>

A;Note: the amino end of the mature protein was determined

C;Comment: Phosphorylation is regulatory and inactivates eEF-2.

C;Genetics:

A;Gene: GDB:EEF2; EF2

A;Cross-references: GDB:119104; OMIM:130610

A;Map position: 19pter-19q12

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; phosphoprotein; protein biosynthesis

F;2-858/Product: translation elongation factor eEF-2 #status predicted <MAT>

F;20-161/Domain: translation elongation factor Tu homology <ETU>

F;26-33/Region: nucleotide-binding motif A (P-loop)

F;158-161/Region: GTP-binding NKXD motif

F;57,59/Binding site: phosphate (Thr) (covalent) (by elongation factor 2 kinase)

#status predicted

F;715/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status predicted

Query Match 34.4%; Score 870; DB 1; Length 858;
Best Local Similarity 37.3%; Pred. No. 2.4e-56;

Matches 182; Conservative 95; Mismatches 189; Indels 22; Gaps 5;

```
Qy      9 NLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTF-----GDI-----NYQEFAKRLW 56
      |::: |:| | | | :   ||| ||::|   | :           :: |:|
Db     202 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261

Qy     57 GDIYFNPKTRKFTKKAPTSSSQ--RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
      || |:|   ||:| | :   :   |:| : ||:|::|:   ::           : ::| |
Db     262 GDRYFDPANGKFSKSATSPEGKKLPRTFCQLILDPIFKVFDAIMNFKKEETAKLIEKLDI 321

Qy    114 HLTKEELKLNIRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDL 173
      | | :   :||:| | ::   : | | :|||   | : | | | | :
Db     322 KLDSEDKDKEGKPLLKAVMRRLWPAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 381

Qy    174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEED 233
      :   ||| |||| : :|| | | :|: |||| || : | |::| |||   :||
Db     382 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKED 441

Qy    234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
      : : | : : || : || || | : |||| :|| |||   | |   | :|
Db     442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498

Qy    294 FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVM 353
      |: : |:::|||   ||::|||:::|::: || | :   :|||::| | |||::: :
Db     499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHTIAGAGELHLEICL 558

Qy    354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
      || : :: | || :|||::: ||| | |:: | :::||| |: : | | |   |||||:
Db     559 KDLEEDHACIPIKKS DPVVS YRETVSEESNVLCLSKSPNKHNRLYMKARFFPDGLAEDID 618

Qy    414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
      |   | :   :   ||:|: || || |||| ||||| | | :   |
Db     619 KGEVSARQELKQRARYLAEKYEWDVAEARKIWCFGPDGTGPNILTDITKGVQ-----YLNE 674

Qy    474 VKDSIVQG 481
      :|||:| |
Db     675 IKDSVVAG 682
```

RESULT 5

EFRT2

translation elongation factor eEF-2 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 28-Feb-1990 #sequence_revision 02-Dec-1994 #text_change 19-Jan-2001

C;Accession: S04007; B25440; A11180; S02526; S62217

R;Oleinikov, A.V.; Jokhadze, G.G.; Alakhov, Y.B.

FEBS Lett. 248, 131-136, 1989

A;Title: Primary structure of rat liver elongation factor 2 deduced from the cDNA sequence.

A;Reference number: S04007; MUID:89252028; PMID:2721670

A;Accession: S04007

A;Molecule type: mRNA

A;Residues: 1-693 <OLE>

A;Cross-references: EMBL:Y07504; NID:g56081

A;Note: the sequence in GenBank entry RNEF2R (PID:g56082) includes the sequence of B25440 not determined by these authors

R;Kohno, K.; Uchida, T.; Ohkubo, H.; Nakanishi, S.; Nakanishi, T.; Fukui, T.; Ohtsuka, E.; Ikehara, M.; Okada, Y.
 Proc. Natl. Acad. Sci. U.S.A. 83, 4978-4982, 1986
 A;Title: Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: homology with GTP-binding proteins.
 A;Reference number: A25440; MUID:86259716; PMID:3014523
 A;Accession: B25440
 A;Molecule type: mRNA
 A;Residues: 516-858 <KOH>
 A;Cross-references: GB:K03502; NID:g203996; PIDN:AAA41106.1; PID:g203997
 A;Note: nucleotide differences with hamster shown
 R;Robinson, E.A.; Henriksen, O.; Maxwell, E.S.
 J. Biol. Chem. 249, 5088-5093, 1974
 A;Title: Elongation factor 2. Amino acid sequence at the site of adenosine diphosphate ribosylation.
 A;Reference number: A11180; MUID:74301260; PMID:4368673
 A;Accession: A11180
 A;Molecule type: protein
 A;Residues: 702-714,'X',716 <ROB>
 A;Experimental source: liver
 R;Nilsson, L.; Nygard, O.
 Eur. J. Biochem. 171, 293-299, 1988
 A;Title: Structural and functional studies of the interaction of the eukaryotic elongation factor EF-2 with GTP and ribosomes.
 A;Reference number: S02526; MUID:88111683; PMID:3338467
 A;Accession: S02526
 A;Molecule type: protein
 A;Residues: 2-14;68-79;572-585 <NIL>
 R;Guillot, D.; Vard, C.; Reboud, J.P.
 Eur. J. Biochem. 236, 149-154, 1996
 A;Title: Photoaffinity labeling of elongation factor-2 with 8-azido derivatives of GTP and ATP.
 A;Reference number: S62217; MUID:96184892; PMID:8617259
 A;Accession: S62217
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-7;240-244;265-272 <GUI>
 C;Comment: Phosphorylation is regulatory and inactivates eEF-2.
 C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology
 C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; phosphoprotein; protein biosynthesis
 F;2-858/Product: translation elongation factor eEF-2 #status predicted <MAT>
 F;20-161/Domain: translation elongation factor Tu homology <ETU>
 F;26-33/Region: nucleotide-binding motif A (P-loop)
 F;158-161/Region: GTP-binding NKXD motif
 F;57,59/Binding site: phosphate (Thr) (covalent) (by elongation factor 2 kinase) #status predicted
 F;715/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status experimental

Query Match 34.3%; Score 868; DB 1; Length 858;
 Best Local Similarity 37.7%; Pred. No. 3.3e-56;
 Matches 184; Conservative 92; Mismatches 190; Indels 22; Gaps 5;

QY 9 NLILSPLLGNVCFSSSQYSICFTLGSAFIYADTF-----GDINYQEFA-----KRLW 56
 |::: |::| | | : ||| |:::| | : | | |::|

A;Molecule type: DNA
A;Residues: 1-15, 'E', 17-302, 'P', 304-440, 'D', 442-716, 'R', 718-830, 'G', 832-858
<NAK>
A;Cross-references: GB:J03200; NID:g191009; PIDN:AAA50386.1; PID:g387049
A;Note: the authors translated the codon GAA for residue 16 as Lys, CCC for residue 303 as Ala, and GGC for residue 831 as Pro
C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology
C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; phosphoprotein; protein biosynthesis
F;20-161/Domain: translation elongation factor Tu homology <ETU>
F;26-33/Region: nucleotide-binding motif A (P-loop)
F;158-161/Region: GTP-binding NKXD motif
F;715/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status predicted

Query Match 34.2%; Score 863; DB 2; Length 858;
Best Local Similarity 37.1%; Pred. No. 7.8e-56;
Matches 181; Conservative 95; Mismatches 190; Indels 22; Gaps 5;

```

QY      9 NLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTF-----GDI-----NYQEFARKLW 56
      |:::|:| | | | : | | | |:::| | : :: |::|
Db     202 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261

QY     57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGDDVDTSLPRTLDELGI 113
      || |::| |::| | : : |::| |:::|: :: : ::| |
Db     262 GDRYFDPANGKFSKANSNPDGKKLPRTFCQLILDPIFKVFDAMNFRKEETAKLIEKLDI 321

QY     114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVDSDL 173
      | | : :|:::| ::: : | |::| | : | | | :
Db     322 KLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 381

QY     174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
      : | | | |:::| :: | | : | | | | : | |:::| | | : |
Db     382 AMGKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKEE 441

QY     234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
      : : | : : | | : | | | | : | | | | | | | | | | | : |
Db     442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498

QY     294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVM 353
      | : : |:::| | | |:::| |:::| | | : : | | | | | | | | | | :
Db     499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICL 558

QY     354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
      | | : : | | | |:::| | | | | : | : | | | | | | | | | | :
Db     559 KDLEEDHACIPIKKS DPVVS YRETVSEESNVLC LSKSPNKHNR LYM KARFPDGLAEDID 618

QY     414 NEVVQITWNRKKLG EFFTQ KYDWDLLAARS IWA FGPDATGPNILVDDTL PSEVDKALLGS 473
      | | : | : |:::| | | | | | | | | | | | | | | | : |
Db     619 KGEVSARQELKARARYLA EKY EWDVAEARKI WCF GPDGTGPNILTDITKGVQ---YLNE 674

QY     474 VKDSIVQG 481
      : | | | : | |
Db     675 IKDSVVAG 682

```

RESULT 7

S05988

translation elongation factor eEF-2 - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 02-Feb-2001

C;Accession: S05988

R;Grinblat, Y.; Brown, N.H.; Kafatos, F.C.

Nucleic Acids Res. 17, 7303-7314, 1989

A;Title: Isolation and characterization of the *Drosophila* translational elongation factor 2 gene.

A;Reference number: S05988; MUID:90016792; PMID:2508059

A;Accession: S05988

A;Molecule type: DNA

A;Residues: 1-844 <GRI>

A;Cross-references: EMBL:X15805; NID:g7918; PIDN:CAA33804.1; PID:g7919

C;Genetics:

A;Gene: FlyBase:Ef2b

A;Cross-references: FlyBase:FBgn0000559

A;Map position: 2L 39E-F

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; phosphoprotein; protein biosynthesis

F;20-165/Domain: translation elongation factor Tu homology <ETU>

F;26-33/Region: nucleotide-binding motif A (P-loop)

F;162-165/Region: GTP-binding NKXD motif

F;609-611/Region: GTP-binding SAK/L motif

F;701/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status predicted

Query Match 33.5%; Score 845.5; DB 2; Length 844;
Best Local Similarity 37.2%; Pred. No. 1.5e-54;
Matches 175; Conservative 103; Mismatches 183; Indels 9; Gaps 4;

```

Qy      12  LSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIFYNPKTRKFTKK 71
      : | |:| | | : ||| |::|:: | |: : |||: :| |:|: :|
Db      208  VDPSKGSVGFSGSLHGWAFTLKQFSEMYSEKF-KIDVVKLMNRLWGENFFNAKTKKW-QK 265

Qy      72  APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLNIRPLLRLV 131
      : ::||| :|:|:|:|: : : : |::|: | |: : : ||: |
Db      266  QKEADNKRSEFCMYILDPIYKVFDAMNYKKEEIGTLEKIGVTLKHEDKDKDGKALLKTV 325

Qy     132  CKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHTT 191
      : : : | |:| | | :| | | : | : | | | | | : :
Db     326  MRTWLPAGEALLQMIAIHLSPVVAQKYRMEMLYEGPHDDEAAIAVKSCDPDGPLMMYIS 385

Qy     192  KMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRWLWISVARYHI 251
      || | | :|: ||| :| : || :::| ||| :|| : | : : ||
Db     386  KMVPTSDKGRFYAFGRVFAGKVATGQKCRIMGPNYTPGKKEDLYEKAIQRTILMMGRYVE 445

Qy     252  EVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE 311
      : ||:| | : ||| :||| ||| :| : :||: : |::| ||| ||::
Db     446  AIEDVPSGNICGLVGVDQFLVKTGTITT---FKDAHNMKVMKFSVSPVVRVAVEPKNPAD 502

Qy     312  LPKMLDGLRKVNKSYPSTLTKEVESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPV 371
      |||::|:|:|: || | : :||| ||| | |||:| : || : : | :| :|||
Db     503  LPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPLKKS DPV 562

```

Qy 372 VTFCE TVVETSS LKCF AETPN KKNKITMIAE PLEKGLAEDIENE VVQITWNRKKLG EFFQ 431
 |:: ||| | | | ::: ||| |:: | | | : | | |||: | | :
 Db 563 VSYRET VSEESDQMCLSKSPNKHNRLLMKALPMPDGLPEDIDNGEVS AKDEFKARARYLS 622
 Qy 432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLG SVKDSIVQG 481
 |||: | : | | || ||| ||| : : | | : | : |||: | |
 Db 623 EKYDYDVTEARKIWC FGP DGTGPNFILDCTKSVQ----YLNEIKDSVVAG 668

RESULT 8

T21362

hypothetical protein F25H5.4 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000

C;Accession: T21362

R;Steward, C.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19412

A;Accession: T21362

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-852 <WIL>

A;Cross-references: EMBL:Z81068; PIDN:CAB02985.1; GSPDB:GN00019; CESP:F25H5.4

A;Experimental source: clone F25H5

C;Genetics:

A;Gene: CESP:F25H5.4

A;Map position: 1

A;Introns: 1/3; 73/2; 191/3; 250/3; 752/1

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

F;20-173/Domain: translation elongation factor Tu homology <ETU>

Query Match 33.1%; Score 837.5; DB 2; Length 852;

Best Local Similarity 36.2%; Pred. No. 6e-54;

Matches 175; Conservative 103; Mismatches 193; Indels 13; Gaps 5;

Qy 2 AYYSTDEN----LILSPLLGNVCFSSSQYSICFTLG SFAKIYADTFGDINYQEF AKRLWG 57
 | | | : : : | : ||| | | : ||| ||: || | | : | |||
 Db 202 ATYGDDDGPMGPIMVDPSIGNVGFSGSLHGWAFTLKQFAEMYAGKFG-VQVDKLMKNLWG 260
 Qy 58 DIYFNP KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVG DVDTSLPRTLDELGIHLTK 117
 | : | : ||: | : | | : | : | : | : : | : : ||| |
 Db 261 DRFFDLKTKKWS-STQTDESKRGFCQFVLDP I FMVFDAMNIKKDKTAALVEKLG I KLAN 319
 Qy 118 EELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAM 177
 : | | : ||: : : : : | : ||| | : | | | : | :
 Db 320 DEKDLEGKPLMKVFMRKWLPAGDTMLQMI AFHLPSPVTAQKYRMEMLYEGPHDDEAAVAI 379
 Qy 178 SDCDPDGPLMCHTTKMFSTHDGVQFHFPGRVLSGTI HAGQPVKVLGENYTLDEEDSQIC 237
 |||: ||| : : || | | : : ||| || : | : : | | : ||
 Db 380 KTCDPNGPLMMYISKMVPTSDKGRFYAFGRVFS GK VATGMKARIQGP NYVP GK KEDLYEK 439
 Qy 238 TVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
 | : | : : | : : | : | : ||| : || || : : | : ||: :
 Db 440 TIQRTILMMGRFIEPIEDIPSGNIAGLVGVDQYLVKGGTITT---YKDAHNM RMVMKFSVS 496

Qy 298 SVIKIAVEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLR 357
 |::||| |::|||::||::| | | : | | | | | : | | | : : | |
 Db 497 PVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIFEESGEHIIAGAGELHLEICLKDLE 556
 Qy 358 KMYSEIDIKVADPVVTFCETVVETSSLKCFAPETPNKKNKITMIAEPLKGLAEDIENEVV 417
 : : | : | : | | | : | | | : | : | : | | | : | : | : | | | |
 Db 557 EDHACIPLKKSDFVVSRETQVSESQICLSKSPNKHNLHCTAQPMPLDGLADDIEGGTV 616
 Qy 418 QITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSKVDS 477
 | : | | : | : | | | | | | | : | : | : | | |
 Db 617 NARDEFKARAKILAEKYEYDVTEARKIWCFCGPDGTGPNLLMDVTKGVQ----YLNEIKDS 672
 Qy 478 IVQG 481
 : | |
 Db 673 VVAG 676

RESULT 9

A40411

translation elongation factor eEF-2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 02-Feb-2001

C;Accession: A40411

R;Ofulue, E.N.; Candido, E.P.M.

DNA Cell Biol. 10, 603-611, 1991

A;Title: Molecular cloning and characterization of the Caenorhabditis elegans elongation factor 2 gene (eft-2).

A;Reference number: A40411; MUID:92029622; PMID:1930695

A;Accession: A40411

A;Molecule type: mRNA

A;Residues: 1-852 <OFU>

A;Cross-references: GB:M86959; NID:g156278; PIDN:AAD03339.1; PID:g156279

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; phosphoprotein; protein biosynthesis

F;20-173/Domain: translation elongation factor Tu homology <ETU>

F;26-33/Region: nucleotide-binding motif A (P-loop)

F;170-173/Region: GTP-binding NKXD motif

F;709/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status predicted

Query Match 32.7%; Score 827.5; DB 2; Length 852;
 Best Local Similarity 36.0%; Pred. No. 3.3e-53;
 Matches 174; Conservative 104; Mismatches 193; Indels 13; Gaps 5;

Qy 2 AYYSTDEN----LILSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFARLWG 57
 | | | : : : | : | | | | : | | | | : | | : | | |
 Db 202 ATYGDDDGPMGPIMVDPSIGNVGFSGSLHGWAFTLKQFAEMYAGKFG-VQVDKLMKNLWG 260
 Qy 58 DIYFNPKTRKFTTKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTK 117
 | : | : | : | : | | : | : | : | : : | : : : | | |
 Db 261 DRFFDLKTKKWS-STQTDSEKRGFCQFVLDPIMVFDAMNKKDKTAALVEKLGKIKLAN 319
 Qy 118 EELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDLGEAM 177
 : | | : | : : : : : | : | | | | : | | | : | :
 Db 320 DEKDLEGKPLMKVFMRKWLPAAGDTMLQMI AFHLPSPVTAQKYRMEMLYEGPHDDEAAVAI 379

Qy 178 SDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQIC 237
 |||:|||| : :|| | | :|: ||| | | : | :| | | :||
 Db 380 KTCDPNGPLMMYISKMVPTSDKGRFYAFGRVFSQKQVATGMKARIQGPNYVPGKKEDLYEK 439
 Qy 238 TVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
 | : | : : | : : :|:| | : ||| :|| ||| :|| | :||: :
 Db 440 TIQRTILMMGRFIEPIEDIPSGNIAGLVGVQYLVKGGTITT---YKDAHNMVRVMKFSVS 496
 Qy 298 SVIKIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLR 357
 |:::|||| ||::|||:::|::: || | : |||||:| | |||:| : ||
 Db 497 PVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIFEESGEHIIAGAGELHLEICLKDLE 556
 Qy 358 KMYSEIDIKVADPVVTFCEVTVETSSLKCFEATPNKKNKITMIAEPLEKGLAEDIENEV 417
 : : : | : | :|||:: ||| | : | :||| |:: |:: |||:| |
 Db 557 EDHACIPLKSDPVVSYRETVQSESNQICLSKSPNKHNRHLCTAQPMPLGLADDIEGGTV 616
 Qy 418 QITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSKV 477
 | : : ||: : || || ||| |||:|:| | : | :|||
 Db 617 SARDEFKARAKYPGEKYEYAVTEARKIWCFGPDGTGPNLLMDVTKGVQ----YLNEIKDS 672
 Qy 478 IVQG 481
 :| |
 Db 673 VVAG 676

RESULT 10

T14579

translation elongation factor eEF-2 - beet

C;Species: Beta vulgaris (beet)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000

C;Accession: T14579

R;Vogel, R.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z18148

A;Accession: T14579

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-843 <VOG>

A;Cross-references: EMBL:Z97178

A;Experimental source: root

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

C;Keywords: GTP binding; phosphoprotein

F;20-161/Domain: translation elongation factor Tu homology <ETU>

Query Match 32.3%; Score 816.5; DB 2; Length 843;

Best Local Similarity 35.7%; Pred. No. 2.1e-52;

Matches 175; Conservative 111; Mismatches 179; Indels 25; Gaps 8;

Qy 9 NLILS----PLLGN-----VCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRL 55
 |::: |||: | ||: : ||| :|||:| | || : : : ||
 Db 186 NVIMATYEDPLLGDVQVYPEKGTVAFSAGLHGWAFTLSNFAKMYASKFG-VDESKMMERL 244
 Qy 56 WGDIIYFNPKTRKF-TKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIH 114
 ||: :|| |::: || : :| :| ||:| ||: :|| : | : :|||
 Db 245 WGENFFDPATKKWTTKNSGNASCKRGFVQFCYEPIKQIIAACMNDQKDKLLAHVTKLGIQ 304

Qy 115 LTKEELKLNIRPLLRVLCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTTGGVDS DLG 174
 : || | |||: | : : : : | : || | : : | : | : |
 Db 305 MKTEEKDL MGRPLMKRVMQTWLPASSALLEMMIHLPSPATAQRYRVENLYEGPMDDVYA 364
 Qy 175 EAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEEDS 234
 | : : |||: ||| : : || | : | ||| : | : | | : : | || : : |
 Db 365 TAIRNCDPEGPLMLYVSKMIPASDKGRFFAFGRVFAGKVSTGMKVRIMGPNYVPGEKKDL 424
 Qy 235 QICTVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKF 294
 : | | | : : | || || | : | : || | | || : | : : | | : ||
 Db 425 YVKNVQRTVIWMGKKQETVEDVPCGNTVALVGLDQYITKNATLTNEK--ESDAHPIRAMKF 483
 Qy 295 NTTSVIKIAVEPVNPSELPMKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMH 354
 : : | : : || : | : ||| : : || : : || | : : ||| || : | ||| : : :
 Db 484 SVSPVVRVAVQCKVASDLPKLVLEGLKRLAKSDPMVVCSEESGEHIIAGAGELHLEICLK 543
 Qy 355 DLRKMY-SEIDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
 || : : | : ||| : | ||| : : | : : || | | : : | | | : ||| | :
 Db 544 DLQDDFMGGA EIIKSDPVVSFRET VLD RSVRTVMSKSPNKHNRLYMEARPMEEGLAE AID 603
 Qy 414 NEVVQITWNRKKLGEFFQTKYDWDLLAARS IWA FGPDATGPNILVDDTLPSEVDKAL---L 471
 : : | : : | | | : | | ||| : ||| : : || : | : |
 Db 604 EGRIGPRDDPKNRSKILAE EYGWDKDLAKKIWC FGPE TTGPNMVVD-----MCKGVQYL 657
 Qy 472 GSVKDSIVQG 481
 : ||| : | |
 Db 658 NEIKDSVVAG 667

RESULT 11

A41778

translation elongation factor eEF-2 - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein D9481.22; protein O3317; protein YDR385w; protein YOR133w; protein YOR3317w

C;Species: *Saccharomyces cerevisiae*

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001

C;Accession: A41778; S60995; S61689; S67018; S61180; S69669; S63872

R;Perentesis, J.P.; Phan, L.D.; Gleason, W.B.; LaPorte, D.C.; Livingston, D.M.; Bodley, J.W.

J. Biol. Chem. 267, 1190-1197, 1992

A;Title: *Saccharomyces cerevisiae* elongation factor 2. Genetic cloning, characterization of expression, and G-domain modeling.

A;Reference number: A41778; MUID:92112760; PMID:1730643

A;Accession: A41778

A;Molecule type: DNA

A;Residues: 1-842 <PER>

A;Cross-references: EMBL:M59369; NID:g415721; PIDN:AAA21646.1; PID:g549849

A;Note: sequence extracted from NCBI backbone (NCBIN:76324, NCBI P:76328)

R;Van Ness, B.G.; Howard, J.B.; Bodley, J.W.

J. Biol. Chem. 255, 10717-10720, 1980

A;Title: ADP-ribosylation of elongation factor 2 by diphtheria toxin. Isolation and properties of the novel ribosyl-amino acid and its hydrolysis products.

A;Reference number: A44720; MUID:81046928; PMID:7000782

A;Contents: annotation; identification and isolation of dipthamide

R;Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.; Zimmermann, J.; Erfle, H.; Paces, V.; Ansorge, W.

submitted to the EMBL Data Library, August 1995
 A;Description: Sequencing of 51 kilobases on the right arm of chromosome XV from *S. cerevisiae* reveals 30 open reading frames.
 A;Reference number: S60983
 A;Accession: S60995
 A;Molecule type: DNA
 A;Residues: 1-842 <WIE>
 A;Cross-references: EMBL:X90518; NID:g1050808; PIDN:CAA62116.1; PID:g1050821
 A;Genetics: EFT1
 R;Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia, A.; Ansorge, W.; Voss, H.
 submitted to the EMBL Data Library, December 1995
 A;Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome XV.
 A;Reference number: S61643
 A;Accession: S61689
 A;Molecule type: DNA
 A;Residues: 1-842 <BEN>
 A;Cross-references: EMBL:X94335; NID:g1262139; PIDN:CAA64052.1; PID:g1164977
 A;Genetics: EFT1
 R;Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansorge, W.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: S66965
 A;Accession: S67018
 A;Molecule type: DNA
 A;Residues: 1-842 <VOS>
 A;Cross-references: EMBL:Z75041; MIPS:YOR133w; NID:g1420341; PIDN:CAA99332.1; PID:g1420342
 A;Experimental source: strain S288C
 A;Genetics: EFT1
 R;Ding, H.
 submitted to the EMBL Data Library, June 1995
 A;Description: The sequence of *S. cerevisiae* cosmid 9481.
 A;Reference number: S61159
 A;Accession: S61180
 A;Molecule type: DNA
 A;Residues: 1-515 <DIN>
 A;Cross-references: EMBL:U28373; NID:g849184; PIDN:AAB64821.1; PID:g849206
 A;Experimental source: strain S288C (AB972)
 A;Genetics: EFT2
 R;Dietrich, F.S.
 submitted to the EMBL Data Library, July 1995
 A;Description: The sequence of *S. cerevisiae* cosmids 9481, 9509, 9926, 9461, and lambda 3641.
 A;Reference number: S69665
 A;Accession: S69669
 A;Molecule type: DNA
 A;Residues: 1-842 <DIE>
 A;Cross-references: EMBL:U32274; NID:g927313; PIDN:AAB64827.1; PID:g927318
 A;Genetics: EFT2
 R;Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.; Zimmermann, J.; Erfle, H.; Paces, V.; Ansorge, W.
 Yeast 12, 281-288, 1996
 A;Title: Sequencing and analysis of 51 kb on the right arm of chromosome XV from *Saccharomyces cerevisiae* reveals 30 open reading frames.
 A;Reference number: S63860; MUID:97060020; PMID:8904341

A;Accession: S63872
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-842 <WIW>
 A;Cross-references: EMBL:X90518; NID:g1050808; PIDN:CAA62116.1; PID:g1050821
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
 C;Genetics: <EFT1>
 A;Gene: SGD:EFT1
 A;Cross-references: MIPS:YOR133w; SGD:S0005659
 A;Map position: 15R
 A;Note: YOR133w
 C;Genetics: <EFT2>
 A;Gene: SGD:EFT2
 A;Cross-references: MIPS:YDR385w; SGD:S0002793
 A;Map position: 4R
 A;Note: YDR385w
 C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology
 C;Keywords: diphthamide; GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F;20-161/Domain: translation elongation factor Tu homology <ETU>
 F;26-33/Region: nucleotide-binding motif A (P-loop)
 F;158-161/Region: GTP-binding NKXD motif
 F;699/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine (His) #status predicted

Query Match 32.0%; Score 807.5; DB 2; Length 842;
 Best Local Similarity 35.7%; Pred. No. 1e-51;
 Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;

Qy	14	PLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFARLWGDYFNPKTRKFTKK--	71
		: : :: : : : :	
Db	204	PARGTVAFGSGLHGWAFTIRQFATRYAKKFG-VDKAKMMDRLWGDSSFFNPKTKKWTNKDT	262
Qy	72	-APTSSSQRSFVEFILEPLYKILAQQVVGDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL	130
		: : ::: : : :: : : ::	
Db	263	DAEGKPLERAFNMFILDPIFRLETAIMNFKKDEIPVLEKLEIVLKGDEKDLGKALLKV	322
Qy	131	VCKKFFGEFTGFVDMCVQHIPSQVGAQPKIEHTYTTGGVSDSLGEAMSDCDPDGPLMCHT	190
		: : : : : :	
Db	323	VMRKFLPAADALLEMIVLHLPSPVTAQAYRAEQLYEGPADDANCIKNCDPKADIMLYV	382
Qy	191	TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRWLISVARYH	250
		: : : : : : :: : : : : :	
Db	383	SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPKKDDLFKAIQRVVLMGREFV	442
Qy	251	IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS	310
		:: : : :: :: :: : : : : : ::: :	
Db	443	EPIDDCPAGNIIGLVGIDQFLKGTGLTT---SETAHNMKVMKFSVSPVQVAVEVKNAN	499
Qy	311	ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADP	370
		: :: : : : : : : : : : : : : : ::	
Db	500	DLPKLVEGLKRLSKSDPCVLTYMSESGEHIVAGTGELHLEICLDLEHDHAGVPLKISPP	559
Qy	371	VVTFCE TVVETSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVQITWNRKKLGEFF	430
		: : :: :: : : :: :	

Db 560 VVAYRETVESESSQTALSKSPNKHNRILKAEPIDEEVSLAIENGIINPRDDFKARARIM 619

Qy 431 QTKYDWDLLAARSIWAFGPDATGENILVDDTLPSEVDKAL--LGSVKDSIV 479
 | ||: || || |||| |||::| | ||: | :|||:|

Db 620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSVV 664

RESULT 12

T41697

translation elongation factor eEF-2 [similarity] - fission yeast
 (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000

C;Accession: T41697; T39256; T38896; T51993; T51994; T42427

R;Wedler, H.; Duesterhoeft, A.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999

A;Reference number: Z22010

A;Accession: T41697

A;Molecule type: DNA

A;Residues: 1-842 <WED>

A;Cross-references: EMBL:AL121859; PIDN:CAB58373.1; GSPDB:GN00068;

SPDB:SPCP31B10.07

A;Experimental source: strain 972h-; clone pl p31B10

A;Genetics: SP2

R;Stevens, K.; Churcher, C.; McDougall, R.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1999

A;Reference number: Z21838

A;Accession: T39256

A;Molecule type: DNA

A;Residues: 31-842 <STE>

A;Cross-references: EMBL:AL109734; PIDN:CAB52147.1; GSPDB:GN00066;

SPDB:SPAPYUK71.04c

A;Experimental source: strain 972h-; cosmid pYUK71

A;Genetics: SP1

R;Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999

A;Reference number: Z21816

A;Accession: T38896

A;Molecule type: DNA

A;Residues: 1-63 <SEE>

A;Cross-references: EMBL:AL122032; PIDN:CAB58724.1; GSPDB:GN00066;

SPDB:SPAC513.01c

A;Experimental source: strain 972h-; cosmid c513

A;Genetics: SP1

R;Mita, K.; Morimyo, M.; Ito, K.; Sugaya, K.; Ebihara, K.; Hongo, E.; Higashi,
 T.; Hirayama, Y.; Nakamura, Y.

Gene 187, 259-266, 1997

A;Title: Comprehensive cloning of Schizosaccharomyces pombe genes encoding
 translation elongation factors.

A;Reference number: Z22377; MUID:97254480; PMID:9099890

A;Accession: T51993

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-820, 'DVG', 824-842 <MIT1>

A;Cross-references: EMBL:D83975; PIDN:BAA23590.1

A;Experimental source: strain JY333, gene EF-2.1

A;Accession: T51994

A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-820,'DVG',824-842 <MIT2>
 A;Cross-references: EMBL:D83976; PIDN:BAA23591.1
 A;Experimental source: strain JY333, gene EF-2.2
 R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A;Title: Identification of open reading frames in Schizosaccharomyces pombe
 cDNAs.
 A;Reference number: Z17323; MUID:98162722; PMID:9501991
 A;Accession: T42427
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 177-322,'D',324-463,'A',465-
 700,'WWWSNYSYCSSCRLRFYPFGFSHHPGARFLGRNPSFRKRYGWYLLCP' <YOS>
 A;Cross-references: EMBL:D89151; NID:g1749509; PIDN:BAA13813.1; PID:g1749510
 A;Experimental source: strain PR745
 C;Genetics: <SP1>
 A;Gene: SPDB:SPAPYUK71.04c; SPDB:SPAC513.01c
 A;Map position: 1
 C;Genetics: <SP2>
 A;Gene: SPDB:SPCP31B10.07
 A;Map position: 3
 C;Superfamily: translation elongation factor 2; translation elongation factor Tu
 homology
 C;Keywords: GTP binding; protein biosynthesis
 F;20-161/Domain: translation elongation factor Tu homology <ETU>

Query Match 31.7%; Score 801.5; DB 2; Length 842;
 Best Local Similarity 36.2%; Pred. No. 2.8e-51;
 Matches 170; Conservative 105; Mismatches 183; Indels 11; Gaps 4;

Qy	14	PLLGNVCFSSSQYSICFTLGFSFAKIYADTFGDINYQEEFAKRLWGDIFYNPKTRKFTKKAP	73
Db	204	PDKGTVAFAASGLHGWAFTVRQFANRYAKKFG-IDRNKMMQRLWGENYFNPKTKKWSKSAT	262
Qy	74	T---SSSQRSFVEFILEPLYKILAQQVVGDVDTSLPRTLDELGIHILTKEELKLNIRPLLRL	130
Db	263	DANGNSNQRAFNMFIIDPIYRIFDAVMNSRKDEVFTLLSKLEVTIKPDEKELEGKALLKV	322
Qy	131	VCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DLGEAMSDCDPDG PLMCHT	190
Db	323	VMRKFLPAADALMEMIVLHLPSPKTAQQYRAETLYEGPMDDECAVGIRNCDANAPLM IYV	382
Qy	191	TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYH	250
Db	383	SKMVPTSDRGRFYAFGRVFSGTVRSGLKVRIQGP NYVP GKDDLFIKAIQRTVLM MGSR I	442
Qy	251	IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS	310
Db	443	EPIEDCPAGNIIGLVGVDQFLVKSGTLTT---SEVAHNMKVMKFSVSPVQVAVEVKNGN	499
Qy	311	ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADP	370
Db	500	DLPKLVEGLKRLSKSDPCVLCTTSES GEHIVAGAGELHLEICLKD LQEDHAGIPLKISPP	559
Qy	371	VVTFCETTVETSSLKCF AETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNRKKLGEFF	430

Db 560 VVSYSRESVSEPSSMTALSKSPNKHNRIFMTAEPMSEELSVAIETGHVNPRDDFKVRARIM 619
 QY 431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSSEVDKALLGSKVKSIV 479
 Db 620 ADEFGWDVTDARKIWCFGPDTTGANVVVDQTKA----VAYLNEIKDSVV 664

RESULT 13

A96602

elongation factor EF-2 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C;Accession: A96602

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;
 Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;
 Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;
 Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;
 Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.;
 Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;
 Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-
 Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu,
 S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda,
 M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.;
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;
 Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken,
 S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.;
 Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A96602

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-846 <STO>

A;Cross-references: GB:AE005173; NID:g6056373; PIDN:AAF02837.1; GSPDB:GN00141

C;Genetics:

A;Gene: T6H22.13

A;Map position: 1

C;Superfamily: translation elongation factor 2; translation elongation factor Tu
 homology

Query Match 31.4%; Score 793.5; DB 2; Length 846;

Best Local Similarity 35.1%; Pred. No. 1.1e-50;

Matches 172; Conservative 115; Mismatches 178; Indels 25; Gaps 8;

QY 9 NLILS----PLLGN-----VCFSSSQYSICFTLGSFAKIYADTFGDINYQEFKRL 55
 |::: |||: | ||: : ||| :|||: || : :||
 Db 189 NVIMATYEDPLLGDVQVYPEKGTVAFSAGLHGWAFTLTNFAKMYASKFGVVE-SKMMERL 247
 QY 56 WGDIYFNPKTRKFT-KKAPTSSSQRSFVEFILEPLYKILAQVVGDDVDTSLPRTLDELGIH 114
 ||: :||: |||: | : : :| ||: ||: :||: : | | :||:
 Db 248 WGENFFDPATRKWSGKNTGSPTCKRGFVQFCYEPIKQIIATCMNDQKDKLWPLAKLGVS 307
 QY 115 LTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDLG 174

```

      :  :| :| :||:: | : :      | :| : |::||      : ::|: | | :|
Db      308 MKNDEKELMGKPLMKRVMQTLWPASTALLEMMIFHLPSHTAQRYRVENLYEGPLDDQYA 367

QY      175 EAMSDCDPDGPILMCHTTKMFSTHDGVQFHFPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDS 234
      |: :|||:|||| : :||      | :|      ||| :| : | |::| || :::|
Db      368 NAIRNCDPNGPLMLYVSKMIPASDKGRFFAFGRVFAGKVSTGMKVRIMGPNYIPGEKKDL 427

QY      235 QICTVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKF 294
      :| | | : :      | || || | : |::| | | ||:| : :| | :||
Db      428 YTKSVQRTVIWMGKRQETVEDVPCGNTVAMVGLDQFITKNATLTNEK-EVDAHPIRAMKF 486

QY      295 NTTSVIKIAVEPVNPSELPRMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDVCMH 354
      : : |::||:      |:|||::|::| || | :      :|||::| | |||:| : :
Db      487 SVSPVVRVAVQCKVASDLPKLVEGLKRLAKSDPMVVCTMEESGEHIVAGAGELHLEICLK 546

QY      355 DLRKMY-SEIDIKVADPVVTFCETVVETSSSLKCFEAETPNKKNKITMIAEPLEKGLAEDIE 413
      ||: :      :| :|||:| ||| : |:      ::||| |:: | | |::||| |:
Db      547 DLQDDFMGGAEIISKSDPVVSFRETVCDRSTRTVMSKSPNKHNRLYMEAREPMEEGLAE AID 606

QY      414 NEVVQITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--L 471
      : :      : | : : ||      |: |||||: |||::||      : | : |
Db      607 DGRIGPRDDPKIRSKILAE EFGWDKDLAKKIWAFGPETTGPNMVVD-----MCKGVQYL 660

QY      472 GSVKDSIVQG 481
      :|||:| |
Db      661 NEIKDSVVAG 670

```

RESULT 14

G90128

elongation factor EF-2 [imported] - Guillardia theta nucleomorph

C;Species: nucleomorph Guillardia theta

A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001

C;Accession: G90128

R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reith, M.; Cavalier-Smith, T.; Maier, U.G.

Nature 410, 1091-1096, 2001

A;Title: The highly reduced genome of an enslaved algal nucleus.

A;Reference number: A99082; MUID:11323671; PMID:11323671

A;Accession: G90128

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-848 <DOU>

A;Cross-references: GB:AF083031; NID:g13794345; PIDN:AAK39722.1; GSPDB:GN00152

C;Genetics:

A;Gene: EF2

A;Map position: 3

A;Genome: nucleomorph

C;Superfamily: translation elongation factor 2; translation elongation factor Tu homology

C;Keywords: nucleomorph

Query Match 31.1%; Score 787; DB 2; Length 848;

Best Local Similarity 34.0%; Pred. No. 3.4e-50;

Matches 174; Conservative 118; Mismatches 176; Indels 44; Gaps 11;

F;26-33/Region: nucleotide-binding motif A (P-loop)
 F;158-161/Region: GTP-binding NKXD motif
 F;702/Modified site: 2'-[3-carboxamido-3-(trimethylammonio)propyl]histidine
 (His) #status predicted

Query Match 30.8%; Score 779; DB 2; Length 845;
 Best Local Similarity 36.0%; Pred. No. 1.3e-49;
 Matches 169; Conservative 103; Mismatches 188; Indels 10; Gaps 6;

```

Qy      14 PLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWGDYFNPKTRKFTKK-A 72
      | | | ||: : ||| || :|| || : : :||| |: |||:|
Db      204 PEAGTVSFSAGLHGWAFTLTVFANMYAAKFG-TDKRMMEKLWGDNFFDATRKWTKKHT 262

Qy      73 PTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIH--LTKEELKLNIRPLLRL 130
      : :| | :|| ||: :: : | | | :| :: | | :| :||:
Db      263 GADTKRGFCQFIYEPIKTVIEAMNDNKDKLFDLLKKLNVYSKLPEDRELMGKPLMKR 322

Qy     131 VCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSDLGEAMSDCDPDGPI MCHT 190
      | : : :|| : |:|| | ::: | | :| | :|| |||| :
Db     323 VMQTLWLP AHEALLEMMIWHLPSPAKAQKYRVDVLYEGPLDDTYATAVRNCDADGPLMMYV 382

Qy     191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTL EDEEDSQICTVGR LWISVARYH 250
      :|| | :|: |||| || | | :|::| || :|| :|| : :|
Db     383 SKMIPAADKGRFYAFGRVFSGRIATGRKVRIMGPNYVPGQKKDLYVKT VQRTVLCMGR RQ 442

Qy     251 IEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
      | || || | : |:|| | | ||:|: : | : :||: : |::||| |
Db     443 EAVEDVPCGNTVALVGLDQFITKNATLTDEK-CEDAHTIKAMKFSVSPVVRVAVEPKVAS 501

Qy     311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKM Y-SEIDIKVAD 369
      :|||:::|::: || | : :||:| || | |||:| : ||: : :||:
Db     502 DLPKLVEGLKRLAKSDPMVQCTI EETGEHIIAGAGELHLEICLKDIQDDFMGGA EIRVSE 561

Qy     370 PVVTFCETVVETSSLKCAETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNRKKLGEF 429
      |||:| |||: || :::|| |:: | | | :| ||| | : : :| :
Db     562 PVVSFRET VIGTSDHVVMKSPNKHNRLYMQARPMEDGLAE AIDEGKIGPRDDPKVRSKI 621

Qy     430 FQTKYDWDLLAARSIWAFGP DATGPNILVDDTLPSEVDKALLGSKV KDSIV 479
      : : || | : | |||| ||||: | | : | :||:|
Db     622 LSEEFGWDKELAKKILAFGPDTTGPNMVT DITKGVQ----YLNEIKDSVV 667

```

Search completed: January 30, 2004, 11:27:00
 Job time : 27.2023 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 11:26:28 ; Search time 54.2763 Seconds
(without alignments)
1841.751 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSKVKSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description

1	870	34.4	856	12	US-10-116-275-272	Sequence 272, App
2	845.5	33.5	844	14	US-10-108-605-185	Sequence 185, App
3	806	31.9	884	15	US-10-128-714-4037	Sequence 4037, Ap
4	743	29.4	845	9	US-09-874-923-110	Sequence 110, App
5	743	29.4	845	10	US-09-991-496-110	Sequence 110, App
6	554.5	21.9	361	9	US-09-874-923-85	Sequence 85, Appl
7	554.5	21.9	361	10	US-09-991-496-85	Sequence 85, Appl
8	431.5	17.1	1068	12	US-10-320-797-3336	Sequence 3336, Ap
9	406.5	16.1	1044	12	US-10-032-585-7630	Sequence 7630, Ap
10	387.5	15.3	1013	15	US-10-128-714-3410	Sequence 3410, Ap
11	379	15.0	1087	15	US-10-128-714-8410	Sequence 8410, Ap
12	325	12.9	543	9	US-09-927-738-2	Sequence 2, Appli
13	188	7.4	693	9	US-09-815-242-5238	Sequence 5238, Ap
14	188	7.4	715	9	US-09-815-242-12443	Sequence 12443, A
15	188	7.4	715	9	US-09-815-242-12736	Sequence 12736, A
16	187.5	7.4	693	9	US-09-815-242-13233	Sequence 13233, A
17	178	7.0	702	9	US-09-815-242-11858	Sequence 11858, A
18	177	7.0	691	9	US-09-815-242-4977	Sequence 4977, Ap
19	177	7.0	693	9	US-09-815-242-10481	Sequence 10481, A
20	177	7.0	700	9	US-09-815-242-11060	Sequence 11060, A
21	169	6.7	704	9	US-09-912-020-248	Sequence 248, App
22	169	6.7	704	9	US-09-815-242-10363	Sequence 10363, A
23	168	6.6	704	9	US-09-815-242-13965	Sequence 13965, A
24	162	6.4	99	10	US-09-925-300-1750	Sequence 1750, Ap
25	161.5	6.4	705	10	US-09-738-626-4055	Sequence 4055, Ap
26	160	6.3	692	9	US-09-815-242-11568	Sequence 11568, A
27	160	6.3	709	15	US-10-156-761-12453	Sequence 12453, A
28	155	6.1	696	15	US-10-156-761-9338	Sequence 9338, Ap
29	148.5	5.9	651	12	US-10-289-762-589	Sequence 589, App
30	141.5	5.6	187	12	US-10-264-049-2383	Sequence 2383, Ap
31	129	5.1	607	9	US-09-815-242-13791	Sequence 13791, A
32	127	5.0	591	9	US-09-815-242-10419	Sequence 10419, A
33	127	5.0	591	12	US-10-287-274-425	Sequence 425, App
34	118.5	4.7	732	15	US-10-156-761-14353	Sequence 14353, A
35	118	4.7	532	9	US-09-815-242-11727	Sequence 11727, A
36	117	4.6	615	12	US-10-104-047-2952	Sequence 2952, Ap
37	111.5	4.4	599	9	US-09-815-242-11326	Sequence 11326, A
38	111.5	4.4	599	12	US-09-882-227-602	Sequence 602, App
39	111.5	4.4	637	10	US-09-738-626-4721	Sequence 4721, Ap
40	110.5	4.4	409	12	US-10-264-237-1882	Sequence 1882, Ap
41	110.5	4.4	752	11	US-09-815-379-6	Sequence 6, Appli
42	109.5	4.3	693	12	US-10-369-493-22490	Sequence 22490, A
43	106	4.2	2244	12	US-10-369-493-2084	Sequence 2084, Ap
44	105.5	4.2	1226	9	US-09-815-242-13646	Sequence 13646, A
45	103.5	4.1	528	12	US-10-032-585-7031	Sequence 7031, Ap

ALIGNMENTS

RESULT 1
 US-10-116-275-272
 ; Sequence 272, Application US/10116275
 ; Publication No. US20030211476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Elan Pharmaceutical Technology

```
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and
Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell
Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-272
```

```
Query Match          34.4%; Score 870; DB 12; Length 856;
Best Local Similarity 37.3%; Pred. No. 5.9e-78;
Matches 182; Conservative 95; Mismatches 189; Indels 22; Gaps 5;
```

```
Qy      9 NLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTF-----GDI-----NYQEFARKLW 56
      |::: |:|| | | | : | | | |:::| | : : |:||
Db      202 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 261

Qy      57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
      || |:| |:| | : : |:| : |:|:::| : : :::| |
Db      262 GDRYFDPANGKFSKSATSPEGKKLPRTFCQLILDPIFKVFDAIMNFKKEETAKLIEKLDI 321

Qy      114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDL 173
      | | : :||:| |::: : | |:|| | : | | | :
Db      322 KLDSEDKDKEGKPLLKAVMRRWLPGDALLQMITIHLPSVTAQKYRCELLYEGPPDDEA 381

Qy      174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYLTLEDEED 233
      : ||| ||| : :| | :|: ||| | : | |:|: ||| :||
Db      382 AMGIKSCDPKGPLMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKED 441

Qy      234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
      : : | : : || : || || | : ||| :||| || | | | :|
Db      442 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 498

Qy      294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVM 353
      |: : |:|: ||| |:|: |||: |||: |||: ||| : |:|: |||: | | |:|: :
Db      499 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICL 558

Qy      354 HDLRKMYSEIDIKVADPVVTF CETVVETSSLKCAETPNKKNKITMIAEPLEKGLAEDIE 413
      || : : | | | :|||: ||| | |:| : :||| |:| | | | | |:|:
Db      559 KDLEEDHACIPIKSDPVVSRET VSEESNVLCLSKSPNKHNRLYMKARFPDGLAEDID 618

Qy      414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKALLGS 473
      | : : | : |:|: || || ||| |||| | | : |
Db      619 KGEVSARQELKQRARYLAEKYEWDVAEARKIWCFGPDGTGPNILTDITKG VQ---YLNE 674

Qy      474 VKDSIVQG 481
```

Db :||||| |
675 IKDSVVAG 682

RESULT 2

US-10-108-605-185

; Sequence 185, Application US/10108605

; Publication No. US20020160934A1

; GENERAL INFORMATION:

; APPLICANT: Broadus, Julie

; APPLICANT: Stam, Lynn

; APPLICANT: Bachmann, Jane

; APPLICANT: Kamdar, Kim

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE

; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

; FILE REFERENCE: 31133B

; CURRENT APPLICATION NUMBER: US/10/108,605

; CURRENT FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: US 09/761,142

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/176,418

; PRIOR FILING DATE: 2000-01-14

; NUMBER OF SEQ ID NOS: 361

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 185

; LENGTH: 844

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-10-108-605-185

Query Match 33.5%; Score 845.5; DB 14; Length 844;

Best Local Similarity 37.2%; Pred. No. 1.7e-75;

Matches 175; Conservative 103; Mismatches 183; Indels 9; Gaps 4;

```
QY      12 LSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFKRLWGDIFYNPKTRKETKK 71
      : | | : | | : | | | | : : | | : : | | | : | | | : |
Db      208 VDPKSGSVGFGSGLHGWAFTLKQFSEMYSEKF-KIDVVKLMNRLWGENFFNAKTKKW-QK 265

QY      72 APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLRLRV 131
      : : | | | : | | : | | : : : : : | | : : : | | : |
Db      266 QKEADNKRSCMYILDPIYKVFDAIMNYKKEEIGTLLEKIGVTLKHEDKDKDGKALLKTV 325

QY     132 CKKFFGEFTGFVDMCVQHIPSQVGAQPKIEHTYTGGVSDLDGEAMSDCDPDGPILMCHTT 191
      : : : : | | : | | | | | : : | | : : | | | | | : :
Db     326 MRTWLPAGEALLQMIHLPSPVVAQKYRMEMLYEGPHDDEAAIAVKSCDPDGPLMMYIS 385

QY     192 KMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYHI 251
      | | | | : | : | | | : | : | | : : | | : | : | : |
Db     386 KMVPTSDKGRFYAFGRVFAGKVATGQKCRIMGPNYTPGKKEDLYEKAIQRTILMMGRYVE 445

QY     252 EVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE 311
      : | | : | : | | : | | | | : : | : : | : : | | : |
Db     446 AIEDVPSGNICGLVGVDQFLVKTGTITT---FKDAHNMKVMKFSVSPVVRVAVEPKNPAD 502

QY     312 LPKMLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADPV 371
      | | : : | | : : | | : : | | | | : | | | : : | | : : |
```

Db 503 LPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPLKKS DPV 562

Qy 372 VTF CETV VETSS LKCF AETPN KKNKITMIAE PLEKGLAEDIENE VVQITWNRKKLG EFFQ 431
 |:: ||| | | | :::: || | | : || ||:: | | :
 Db 563 VSYRET VSEESDQMCLSKSPNKHNRLLMKALPMPDGLPEDIDNGEVS AKDEFKARARYLS 622

Qy 432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 ||::|: || || |||| |||| ::| | : | :||::| |
 Db 623 EKYDYDVTEARKIWCFGPDGTGPNFILDCTKSVQ----YLNEIKDSVVAG 668

RESULT 3

US-10-128-714-4037

; Sequence 4037, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; TITLE OF INVENTION: Methods of Use

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4037

; LENGTH: 884

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus

US-10-128-714-4037

Query Match 31.9%; Score 806; DB 15; Length 884;

Best Local Similarity 36.0%; Pred. No. 1.7e-71;

Matches 169; Conservative 104; Mismatches 183; Indels 14; Gaps 5;

Qy 14 PLLGNVCFSSSQYSICFTLG SFAKIYADTFGDINYQEFAKRLWGD IYFNP KTRKFTKKAP 73
 | | | | | : ||: || || || :::: :||| | ||::|:|:|
 Db 201 PDRGTVAFGSGLHGWAFTVRQFAVKYAKKFG-VDRKKMLERLWGDNYFN PQTKKWKTKG- 258

Qy 74 TSSSQRSFVEFILEPLYKILAQVVG DVDTSLPRTLDELGIHLTK EELKLNIRPLLR LVCK 133
 ||:| |||:|:| || | | : : ::| | | :| | : ||::: :
 Db 259 -EPEQRAFNMFI LDPIFKIFA AVNNDKTEEIHKLV EKLEIKLASDEKDLKGKALLKVIMR 317

Qy	134	KFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHTTKM	193
Db	318	: : : : : : : :	
Qy	194	KFLPAADAMLEMICIHLPSPVTAQKYRAETLYEGPMDDECAIGIRDCDPKAPLMPLYVSKM	377
Db	378	VPTSDKGRFYAFGRVFSGIVKSGLKVRIQGPNIYIPGKKDDLFVKAIQRTILMMGRFVEPI	437
Qy	254	NRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSELP	313
Db	438	: : : : : : : : : : :	
Qy	314	EDVPAGNIVGLVGVDQFLLKSGTLTT---SETAHNLKVMKFSVSPVVQRGVEVKNAQDLP	494
Db	495	KLVEGLKRLSKSDPCVLTMINESGQHIVAGAGELHLEICLKDLEEDHAGVPLRISDPVVS	554
Qy	374	FCETVVETSSLKCFEAETPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRKKLGEFFQTK	433
Db	555	: : : :: : :: : : : : :	
Qy	434	YRETVGGESSMTALSKSPNKHNRLYVTAQPLGEEVSLAIESGKINPRDDFKARARLLADD	614
Db	615	YDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIVQG	481
Qy	481	: :: : :	
Db	658	YGWDVTDARKIWCFGPDTTGANLLVDOT-----KAVOYLNEIKDSFVSG	658

US-09-874-923-110

; Sequence 110, Application US/09874923

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Bhatia, Ajay

; APPLICANT: Coler, Rhea

; APPLICANT: Probst, Peter

; APPLICANT: Brannon, Mark

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

; FILE REFERENCE: 210121.420C8

; CURRENT APPLICATION NUMBER: US/09/874,923

; CURRENT FILING DATE: 2001-06-04

; NUMBER OF SEQ ID NOS: 122

```
; SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEO ID NO 110

; LENGTH: 845

; TYPE: PRT

; ORGANISM: *Leishmania major*

; FEATURE:

; NAME/KEY: VARIANT

```

; LOCATION: (1) ... (845)

```

```

; OTHER INFORMATION: Xaa = Any Amino Acid

```

US-09-874-923-110

Query Match 29.4%; Score 743; DB 9; Length 845;
 Best Local Similarity 34.7%; Pred. No. 3.5e-65;
 Matches 166; Conservative 102; Mismatches 189; Indels 22; Gaps 9;

```

Qy      12 LSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFARLWGDYFNPKTRKFTKK 71
      :|| | | | : |:|| ||:|| ||:: : :||| | :|: | :|: | :
Db      200 VSPEKGTVAIGSGLQAWAFSLTRFANMYAAKFG-VDELKMRERLWGDNFFDAKNKKWIKQ 258

Qy      72 APTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLL 128
      : : | :| :| :|:|:| | : : : : | | : || | : ||
Db      259 ETNADGERVRAFCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTILTAEEREQVPXKLL 318

Qy     129 RLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTTGGVDS-----LGEAMSDCDPDG 184
      : | | | : | | | :||| : | | : | : : | : : |||
Db     319 KTVMMXFLPAAETLLQMIVAHLSPKKAQAYRAEMLYSGEASPEDKYFMG--IKNCDPAA 376

Qy     185 PLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED-SQICTVGRLW 243
      ||| : :|| | | : | ||| : || : ||| :||| : || : || : ||
Db     377 PLMLYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDKPVQORSV 436

Qy     244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTSVIKIA 303
      : : || | : | || | : ||| : ||| :||| : | | : || : : | : : |
Db     437 LMMGRYQEAVEDMPCGNVVLGVGVDKYIVKSATITD--DGESPHPLRDMKYSVSPVVRVA 494

Qy     304 VEPVNPSELPKMLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDLCVMHDLRMY-SE 362
      || ||| :||| :||| :|| : || : ||| :||| : ||| : || : || : ||
Db     495 VEAKNPSDLPKLVEGLKRLAKSDPLVCSIEESGEHIVAGAGELHLEICLKDLQEDFMNG 554

Qy     363 IDIKVADPVVTFCEVTVETSSSLKCFATPNKKNKITMIAEPLKGLAEDIENEVVQITWN 422
      :| : :||| : ||| : || : | : : || | : || : || : | :
Db     555 APLKISEPVVSFRETVDVSSQQCLSKSANKHNRLFCRGAPLTEELALAMEEGTAGPEAD 614

Qy     423 RKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
      | | | :||| : || | : ||| ||| :|| | | : : : ||| |
Db     615 PKVRARFLADNYEWDVQEARKIWCYGPDRGNPVVVD-----VTKGVQNMAEMKDSEV 667

```

RESULT 5

US-09-991-496-110

; Sequence 110, Application US/09991496

; Patent No. US20020169285A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Webb, John R.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Bhatia, Ajay

; APPLICANT: Coler, Rhea

; APPLICANT: Probst, Peter

; APPLICANT: Brannon, Mark

; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

; FILE REFERENCE: 210121.420C9

; CURRENT APPLICATION NUMBER: US/09/991,496

; CURRENT FILING DATE: 2001-11-20

; NUMBER OF SEQ ID NOS: 137


```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 845
; TYPE: PRT
; ORGANISM: Leishmania major
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 315, 324
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-991-496-110
```

```
Query Match          29.4%; Score 743; DB 10; Length 845;
Best Local Similarity 34.7%; Pred. No. 3.5e-65;
Matches 166; Conservative 102; Mismatches 189; Indels 22; Gaps 9;
```

```
Qy      12 LSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFAKRLWGDIFYFNPKTRKFTKK 71
      :|| | | | : | :| || :|| || :: : :||| | :| :| :| :|
Db      200 VSPEKGTVAIGSGLQAWAFSLTRFANMYAAKFG-VDELKMRERLWGDNFFDAKNKKWIKQ 258

Qy      72 APTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLL 128
      : : | :| :| | :| :| :| :| :| :| :| :| :| :| :|
Db      259 ETNADGERVRR AFCQFCLDPIYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQVPXKLL 318

Qy      129 RLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS D---LGEAMSDCDPDG 184
      : | | | : | :| :| :| :| :| :| :| :| :| :| :|
Db      319 KTVMMXFLPAAETLLQMI VAHLPSPKKAQAYRAEMLYSGEASPEDKYFMG--IKNCDPAA 376

Qy      185 PLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTL EDEED-SQICTVGRLW 243
      ||| : :| | | :| ||| : || : :|| :|| :| || :| :| :|
Db      377 PLMLYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDKPVQRSV 436

Qy      244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTSVIKIA 303
      : : || | :| || | : ||| : ||| :||| : | | :| : :| :| :|
Db      437 LMMGRYQEAVEDMPCGNVVGVLGVGD KYIVKSATITD--DGESPHPLRDMKYSVSPVVRVA 494

Qy      304 VEPVNPSELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLRMY-SE 362
      || ||| :||| : :||| : || | : :||| :||| : ||| :| :| :| :|
Db      495 VEAKNP SDLPKLVEGLKRLAKSDPLVCSIEESGEHIVAGAGELHLEICLKD LQEDFMNG 554

Qy      363 IDIKVADPVVTFCE TVVETSS LKCF AETPNKKNKITMIAEPLEKGLAEDIENEVVQITWN 422
      :| : :||| :| ||| : || :| : :|| :|| :| :| :| :| :|
Db      555 APLKISEPVVSFRET VTDVSSQQCLSKSANKHNRLFCRGAPLTEELALAMEEGTAGPEAD 614

Qy      423 RKKLG EFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKAL--LGSVKDSIV 479
      | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      615 PKVRARFLADNYEWDVQEARKIWCYGP DNRGPNVVVD-----VTKG VQNMAEMKDSFV 667
```

RESULT 6

```
US-09-874-923-85
; Sequence 85, Application US/09874923
; Patent No. US20020081320A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
```



```
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Leishmania major
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 63, 72, 347, 352, 353, 360
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-991-496-85
```

```
Query Match          21.9%; Score 554.5; DB 10; Length 361;
Best Local Similarity 34.6%; Pred. No. 8.1e-47;
Matches 122; Conservative 82; Mismatches 136; Indels 13; Gaps 6;
```

```
Qy      66 RKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKL 122
      :|: |:  : :  |:| :| |:|:|:|  |: :  : : | | : || |:
Db      1  KKWIKQETNADGERVRR AFCQFC LDP IYQIFDAVMNEKKDKVDKMLKSLHVTLTAEEREQ 60

Qy     123 NIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSD----LGEAMS 178
      ||: |  |  : | | |:| | |  : | | |:|  : : | :
Db      61 VPXKLLKTVMMXFLPAAETLLQMI VAHLPS PKKAQAYRAEMLYSGEASPEDKYFMG--IK 118

Qy     179 DCDPDGPI MCHTTKMFSTHDGVQFHPFGRVLSGTI HAGQPVKVLGENY TLEDEED-SQIC 237
      :||| ||| : :|| | | :| ||| : || : ||| :||| :||| :
Db     119 NCDPAAPLMLYISKMVPTADRGRFFAFGRIFSGKVRSGQKVRIMGNNYVYGKKQDLYEDK 178

Qy     238 TVGRLWISVARYHIEVNRVPAGNWNVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
      | | : : ||  | :| || | : ||| : ||| : ||| : |  | :|: :
Db     179 PVQRSVLMMGRYQEAVEDMPCGNVVGVLGVGD KYIVKSATITD--DGESPHPLRDMKYSVS 236

Qy     298 SVIKIAVEPVNPSEL PKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLD CVMHDLR 357
      |::|||  |||:||||:|:|:|:| | | :  : |||||:|:| |||:|: : ||:
Db     237 PVVRVAVEAKNPSDLPKLVEGLKRLAKSDPLVVC SIEESGEHIVAGAGELHLEICLKDLQ 296

Qy     358 KMY-SEIDIKVADPVVTF CETVVETSSLKCF AETPNKKNKITMIAE PLEKGLA 409
      : : :  :|:|:| ||| : || :| : : | | :| : | |
Db     297 EDFMNGAPLKISEPVVSFRET VTDVSSQQCLSKSANKHNRLFCRGAPLTEXLA 349
```

```
RESULT 8
US-10-320-797-3336
; Sequence 3336, Application US/10320797
```

```

; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS
NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3336
; LENGTH: 1068
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3336

```

```

Query Match          17.1%; Score 431.5; DB 12; Length 1068;
Best Local Similarity 22.4%; Pred. No. 1.2e-33;
Matches 135; Conservative 114; Mismatches 191; Indels 163; Gaps 17;

```

```

Qy      7 DENLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDIFYNPKTR 66
      ||:: :| ||| |:|: | || ||::||: | : |||| | :||:
Db      224 DEDIYFAPDRGNVLFASAIIDGWAFLRGKFARLYAEKL-KIKEGNLRRVLWGDWYLDPKTK 282

Qy      67 KFT--KKAPTSSSQRSFVEFILEPLYKILAQVVGVDV-TSLPRTLDELGIHLTKHEELK-L 122
      : || : : ||:|:| :::: |: : : : : : | | :| :|:
Db      283 RVVGRKKLAGRNKLPFMFVQFVLENIWRVYDVTVLNEYNPDAVQKIVTALNIRITPRDLRSK 342

Qy      123 NIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEH-----TYTGGV--D 170
      : | || |: :: | :: || | : : : : : : :| :| :
Db      343 DTRNLNLIMQQWLPLSTATFQSIIEVIPPPPSAQAIRLPYMLHPEKAKAAAAASGGLKAE 402

Qy      171 SDLGEAMSDCD--PDGPLMCHTTKMFSTHDG--VQFHP----- 204
      ::| : :|| : : : :||: | : : |
Db      403 NELERGLYECDQGEAEVVAYVSKMFAVRKGDLPYKPKEMTAEEMRARGREERERRAAL 462

Qy      205 -----FGRVLSGT 212
      | | : | |
Db      463 VAERQAKGEGLDGQPLPEDLAKPLESLSLENIQPATSEKPAVDDSDSEVLLGFSRIFSST 522

Qy      213 IHAGQPVKVLGENYTLDEEDSQI-----CTVGRWLWISVARYHIEVNRVPAG 259
      :| | : : : || : | : : | : | :| :| |
Db      523 LHRGTSLLAILPKF-----DSSLPPSHPHNIKHTVPIIASDLYMMMGRELVSVDSPAG 576

Qy      260 NWVLIEGVDQPIVKTATITEPR-----GNEEAQIFRPLKENTTSVIKIAVEPVNPSE 311
      : | |:: : : :||: | | : : : :|:| | ||:
Db      577 HVCAIGGLNRAVPRSATLWAPDAKGVEEGFGKEALVNLAGVGVGANAIVRVALEPENPSD 636

Qy      312 LPKMLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPV 371
      :||: ||| :||: | :||| ||||: |||:| : ||: : : | : : | :
Db      637 MPKLIRGLRIINQADPCAIEYFVQESGEHVIITAGELHLERCLKDLRERFAKCPQQSAPI 696

Qy      372 VTFCETVVETSSLKCFEAETPNKKN-----KITMIAEPLEKGLAEDIENE 415

```

```

      | | | | : : | | | : | | : : :
Db      697 VPFRETAVKAPDM-----APPKTTGAPRGTINGTVINGLVKFRRLRAMPLPEGVETFLLSQ 751

QY      416 VVQIT-----WNRKKLGEFFQTKY-----DWDLLAARSIWAFGPDATGPNI 456
      | : | : | | : | | : | : | | | : | :
Db      752 QGAISKMLVRERDGEARQLSPEEFWTELERLLNKAGGDW-AGAADRVSFSGPKRVGANL 810

QY      457 LVD 459
      | : |
Db      811 LLD 813

```

RESULT 9

US-10-032-585-7630

; Sequence 7630, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7630

; LENGTH: 1044

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7630

```

Query Match          16.1%; Score 406.5; DB 12; Length 1044;
Best Local Similarity 23.9%; Pred. No. 3.6e-31;
Matches 146; Conservative 102; Mismatches 199; Indels 165; Gaps 22;

```

```

QY      6 TDENLILSPLLGNVCFSSSQYSICFTIGSFAGKIYADTFGDINYQEFKRLWGDIFYNPKT 65
      :||:| :| || |:|: |:::|||||: | :| :| |||| | :|
Db      214 SDEDLYFTPEKNNVIFASAIIDGWAFSVNTFAKIYSKKLG-FSQQALSCTLWGDFYLDKMN 272

QY      66 RKFT--KKAPTSSSQRS--FVEFILEPLYKILAQ-VVGDDVDTSLPRTLDELGIHLTKEEL 120
      :| || :|: || ||: :: : | : : :|| :| :|
Db      273 KKIIPGKKLKNNSNSLKPLFVSLILDQVWAVYENCVIERNQDKLEKIEKLGAKITPRDL 332

QY      121 K-LNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKI-----EHTYTGGVDS--- 172
      : : : || |: : : : : : : :||| | : :| :| |||:
Db      333 RSKDYKNLLNLIMSQWIPLSHAILGSVIEYLPSPIVAQRERIDKILDETIYSAYDSESDK 392

QY      173 -----LGEAMSDCD---PDGPLMCHTTKMFS-----THDGVQ----- 201
      :|| :|| |: : : :| | | :| :|
Db      393 SKLVDPSFVKAMQECDSSHPEHTIAYVSKLLSIPNEDLPKASNAATGGGLTADEIQERGR 452

QY      202 -----FHP----- 204
      |
Db      453 IARELAKKASEAAALAQEGSKNEDEFAIKPKKDPFEWEFEEDDFENEDESDANAVEEST 512

```

QY 205 -----FGRVLSGTIHAGQPVKVLGENYTLDEEDSQI-----CTVGRLWISVARYHI 251
 | | : | | : | | : | | : | | : | | : | | :
 Db 513 ETIVGFTRIYSGSLSRGQKLTVIGPKYDPSLPRDHQTNFEQITNEVEIKDLFLIMGRELIV 572

QY 252 EVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTS-----VIKIAV 304
 : : | | | | | : : | | : : | | | | : : | : | : | | | |
 Db 573 RMEKVPAGNIVGVVGLDNAVLKNATICSPLPEDKPYI----NLASTSTLIHNKPIMKIAV 628

QY 305 EPVNPSELPMKMLDGLRKVNKSYPSLTKV-EESGEHVILGTGELYLDCVMHDLRKMYSE- 362
 | | | | : | : | | : | : | | : | | | : | | : | | : : :
 Db 629 EPTNPIKLAKLERGLDLLAKADPVLEWYVDDDESSELIVCVAGELHLERCLKDLEERFAKG 688

QY 363 IDIKVADPVVTFCEVTVETSSSLKCFATPNKKN-----KITMIAEPLKKG 407
 : : | : | : | | : : | | : : | | : : | | : | |
 Db 689 CEVTVKEPVI PFREGLADD---KISTNTNNNNDDNEDHELDENEDELADLEFDISPLPLE 745

QY 408 LAED-IENE--VVQITWNRKKLGEF-----FQTKYDWDLLAAR-----SIWAFG 448
 : : | | | : : | : | : | | | | : : | : | |
 Db 746 VTQFLIENETIIAEIVNNKQDTHEIRNDFIEKFATIIDNSNLATQFPDTKSFINNIICFG 805

QY 449 PDATGPNILVDD 460
 | | | | : : |
 Db 806 PKRVGPNIFIED 817

RESULT 10

US-10-128-714-3410

; Sequence 3410, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; TITLE OF INVENTION: Methods of Use

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3410

; LENGTH: 1013

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus
US-10-128-714-3410

Query Match 15.3%; Score 387.5; DB 15; Length 1013;
Best Local Similarity 23.2%; Pred. No. 2.8e-29;
Matches 136; Conservative 97; Mismatches 213; Indels 141; Gaps 16;

```
Qy      7 DENLILSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFAKRLWGDIFYNPKTR 66
      || : :| || | |: ||: || :| | | | | :|||:
Db      222 DEEIYFAPEKNNVIFCSAIDGWAFTVRQFAALYERKLG-IKRSILEKVLWGDFYLDPKTK 280

Qy      67 KF--TKKAPTSSSQRSFVEFILEPLY-----KILAQVVGVDVDTSLP----RTLDELGI 113
      : | : : ||: ||: : : :| : : : :| | : | |
Db      281 RVLGPKHLKGRALKPMFVQLVLDSIWAAYEATTALLEKITKSLNITIPPYVLRSDPRNI 340

Qy      114 HLTKEELKLNIRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGG----- 168
      :| : | || | : : :||| : : | |
Db      341 MMTLFSMWL---PL-----STAVLVSVIEYLPSPAAQATRLPGLIEGSPGAGF 386

Qy      169 VSDSLGEAM--SDCDPDGPLMCHTTKMF----- 194
      || : ||| | | : : :||
Db      387 VDKKVKEAMVAFKTGTDAFVVAIVSKMVAIPESELLSSKKRSGATLSADEAREIARKKRE 446

Qy      195 -----STHDG-VQFHP-----FGRVLSGTI 213
      : || | | | | : |||:
Db      447 EIAKMQAEAGNGNEMDDYARVTSAFEVTTLDDGEEQGEPEDEKEDPEHLVGFARLYSGTL 506

Qy      214 HAGQPVKVLGENYTLED---EEDSQICTVGRWLISVARYHIEVNRVPAGNWWLIEGVDQP 270
      | | || : : | : | || | : : | : ||| | | :
Db      507 SVGDEVYVLAPKFSPEHPHAHPEPQKVTVTDLYLLMGRSLEPLKTVPAGVIFGIGGLAGH 566

Qy      271 IVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPISLT 330
      || | | : | : : | : : : :|| | : || : || : :| |
Db      567 ILKNGTLCSQL--EGSINLAGVSLNAPPIVRVSLPANPADLNKMTGLRLLEQSDPCAQ 624

Qy      331 TKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADPVVTF CETV VETSSLK----- 385
      :| ||||| | |||: : : |||: : : :| : : ||: | :
Db      625 YEVLPSGEHVILTAGEHLHERCIKDLRERFAKCEIQTGQTIVPYRETIVSAPEMAPPKP 684

Qy      386 -----CFAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGEFFQTK----- 433
      | :||: : : || : : : | | | : : :||
Db      685 DLGRGCVLAVSPSKQLTVKLRVPLPEAVTDFISKNVGTIKRLQSEKRSAGETKSDAKAS 744

Qy      434 ---YDWDLLAARSIW-----AFGPDATGPNILVDDTLPSEVDKA 469
      : : : : :| ||| ||||| | : : : |
Db      745 NGIFNEEVKEGKELWANVVDRIATFGPRRVGPNILVDATVVNTCENA 791
```

RESULT 11

US-10-128-714-8410

; Sequence 8410, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

```
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus
fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8410
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8410
```

```
Query Match          15.0%; Score 379; DB 15; Length 1087;
Best Local Similarity 21.8%; Pred. No. 2.3e-28;
Matches 139; Conservative 98; Mismatches 218; Indels 184; Gaps 16;
```

```
Qy      7 DENLILSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFARLWGDIFYFNPKTR 66
      || : :| || | : || : || :| | | | | | : || :
Db      242 DEEIYFAPEKNNVIFCSAIDGWAFTVRQFAALYERKLG-IKRSILEKVLWGDFYLDPKTK 300

Qy      67 KF--TKKAPTSSSQRSFVEFILEPLY-----KILAQVVGDDVDTSLP--- 105
      : | : : || : :| : : :| : : : :|
Db      301 RVLGPKHLKGRALKPMFVQLVLD SIWAAYEATGGGKGKGDPALEKITKSLNITIPPYV 360

Qy      106 -RTLDELGIHLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHT 164
      | : | | :| : | || | : : : :| : :
Db      361 LRSRDPRNIMMTLFSMWL---PL-----STAVLVSVIEYLPSPPAQAATRLPGL 406

Qy      165 YTGG-----VSDSLGEAM--SDCDPDGPLMCHTTKMF----- 194
      | | : || | | : : : :|
Db      407 IEGSPGAGFVDKKVKEAMVAFKTGTDAPVVAYVSKMVAIPESELLSSKKRSGATLSADEA 466

Qy      195 -----STHDG-VQFHP----- 204
      : || | |
Db      467 REIARKKREEIAKMQAEAGNGNEMDDYARVTSAFEVTTLDDGEEQGEPEDEKEDPEHLVG 526

Qy      205 FGRVLSGTIHAGQPVKVLGENYTLED---EEDSQICTVGRWLWISVARYHIEVNRVPAGNW 261
      | :| :|| : | || : : | : | | : : : | : |||
Db      527 FARLYSGTSLVSGDEVYVLAPKFSPEHPHAHPEPQKVTVTDLYLLMGRSLEPLKTVPAAGVI 586

Qy      262 VLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSELPMKLDGLRK 321
      | :| :| : | : : : :| : : : :| : |||
Db      587 FGIGGLAGHILKNGTLCSQL--EGSINLAGVSLNAPPIVRVSLPANPADLNKMTGLRL 644
```



```

Qy      322 VNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADPVVTFCTVTVET 381
      : : | | : | | | | | | | | | : : : | | : : : | | : | | : |
Db      645 LEQSDPCAQYEVLPSEGEHVILTAGEHLHLERICIKDLRERFAKCEIQTGQTIVPYRETIVSA 704

Qy      382 SSLK-----CFAETPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRKKLGFEFF 430
      : | : | : : : | | : : : | | | : :
Db      705 PEMAPPKKPDGLRGCVLAVSPSKQLTVKLRVPLPEAVTDFISKNVGTIKRLQSEKRSAG 764

Qy      431 QTK-----YDWDLLAARSIW-- 445
      : | | : : : : : : : : : :
Db      765 ETKSDAKASNGSLETTQQAESGDASGEAREGSQSLSEDFKKELTRIFNEEVKEGKELWAN 824

Qy      446 -----AFGPDATGPNILVDDTLPSEVDKALLGSVKDSI 478
      | | | | | | | | | : : : | | | :
Db      825 VVDRITAFGPRRVGPNILVDATVVNTCEKCLLDDPKQQL 863

```

RESULT 12

US-09-927-738-2

; Sequence 2, Application US/09927738

; Patent No. US20020076799A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongwen

; TITLE OF INVENTION: Compositions and Methods of modulating TGF-B Signaling

; FILE REFERENCE: 17633/1082

; CURRENT APPLICATION NUMBER: US/09/927,738

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/119786

; PRIOR FILING DATE: 1999-02-11

; PRIOR APPLICATION NUMBER: PCT/US00/03561

; PRIOR FILING DATE: 2000-02-11

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 543

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Clone S1 + 28 protein

US-09-927-738-2

Query Match 12.9%; Score 325; DB 9; Length 543;

Best Local Similarity 25.1%; Pred. No. 2e-23;

Matches 85; Conservative 66; Mismatches 103; Indels 84; Gaps 7;

```

Qy      202 FHPFGRVLSGTIHAGQPVKVLGENYT-----LEDEEDSQICTVGRWLW 243
      | | | | | | | : : | | | : | : | :
Db      88 FIAFARVFSGVARRGKKIFVLGPKYSPLEFLRRVPLCFSAPPDGLPQVPHMAYCALENLY 147

Qy      244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKENTTSVIKIA 303
      : : | : | | | : | : : : | | | | : : : |
Db      148 LLMGRELEYLEEVPNGVNLGIGGLQDFVLKSATLC---SLPSCPPFIPLNFEATPIVRVA 204

Qy      304 VEPVNPSELPMKLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLRKMYSI 363
      | | | : | | | : : : | : : : | | | : | : : : |
Db      205 VEPKHPSEMPQLVKGMKLLNQADPCVQILIQETGEHVILVTAGEVHLQRCCLDDLKERFAKI 264

```

QY 364 DIKVADPVVTFCEVTVETSSLKCFAE----- 389
 | |::|:: | ||: : : |
 Db 265 HISVSEPIIPFRETITKPPKVD MVNEEIGKQQKVAVIHQMKEQSKIPEGIQVDS DGLIT 324
 QY 390 --TPNKKNKITMIAEPLKGLAE-----DIENEVVQITWN-----RKKLG 427
 |||| ::: | || : : : | : : | : | : : : | :
 Db 325 ITTPNKLATLSVRAMPLPEEVTQILENSDLIRSMEQLTSSLNEGENTHMIHQKTQEKIW 384
 QY 428 EFFQTKYDWDLLAAR-----SIWAFGPDATGPNILVD 459
 | | : | : | | : ||: || | ||||| :
 Db 385 E-FKGKLEQHLTGRWRNIVDQIWSFGPRKCGPNILVN 421

RESULT 13

US-09-815-242-5238

; Sequence 5238, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 5238

; LENGTH: 693

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5238

Query Match 7.4%; Score 188; DB 9; Length 693;

Best Local Similarity 22.8%; Pred. No. 1.8e-09;

Matches 96; Conservative 70; Mismatches 161; Indels 94; Gaps 16;


```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12443
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12443
```

```
Query Match          7.4%; Score 188; DB 9; Length 715;
Best Local Similarity 22.8%; Pred. No. 1.9e-09;
Matches 96; Conservative 70; Mismatches 161; Indels 94; Gaps 16;
```

```
QY      77 SQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKLNIRPLLRLVCKKFF 136
      :: | : | : | :: : : | : | : || | | | : | :
Db      232 ARASLIEAVAETSDELMEKYLGDEEISV-----SELKEAIRQATTNV--EFY 276

QY      137 GEFTG-----FVDMCVQHPSPKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPL 186
      | : : : | : | | | : : | : | : | : |
Db      277 PVLCGTAFKNKGVLMLDAVIDYLPSP-LDVKPIIGHRASNPPE-----EVIKADDSAEF 331

QY      187 MCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSDQICTVGRWLISV 246
      | : | | | | | | | : | | | | | | | | |
Db      332 AALAFKVM-TDPYVGKLTFFRVYSGTMTSGSYVK-----NSTKGKRE-----RVGRLLQMH 381

QY      247 ARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEP 306
      | | : | : | : | : | | | : | : | : | : |
Db      382 ANSRQEIDTVYSGDIAAAVGLKD-----TGTGDTLCGEKNDIILESMEF-PEPVIHLSVEP 436

QY      307 VNPSELPMKLDGLRKVNKSYPSLTTKV-EESGEHVILGTGELYLDCVMHDLRKMYSEIDI 365
      : : | | | : : | : | : | : | : | : | : :
Db      437 KSKADQDKMTQALVKLQEEDPTFHAHTDEETGQVIIGGMGELHLDILVDRMKKEFN-VEC 495

QY      366 KVADPVVTFCEVTVVETSSLKCFaETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKK 425
      | | : : | | | | : | : : : : | : |
Db      496 NVGAPMVSURET-----FKSSAQVQGFESR-----QSGGRGQYGDVHIE----- 534

QY      426 LGEFFQTKYDWDLLAARSIAWAFGPDATG-----PNILVDDTLPSEVDKALLGSVKDSIVQ 480
      | | : | | | : | : | : | : | : | :
Db      535 -----FTPNETGAGFEFENAIVGGVVPREYIPSVEAGLKDAMEN 573

QY      481 G 481
      |
Db      574 G 574
```

```
RESULT 15
US-09-815-242-12736
; Sequence 12736, Application US/09815242
```


Qy 307 VNPSELPKMLDGLRKVNKSYPSLTTKV-EESGEHVILGTGELYLDCVMHDLRKMYSEIDI 365
 : :: || | | : : | : ||:| : | | ||:| | : : : : : :
 Db 437 KSKADQDKMTQALVKLQEEDPTFHAHTDEETGQVIIGMGELHLDILVDRMKKEFN-VEC 495
 Qy 366 KVADPVVTFCETVVETSSLKCFEAETPNKKNKITMIAEPELEKGLAEDIENEVVQITWNRKK 425
 | | : : : | | | : | : : : : | : |
 Db 496 NVGAPMVSURET-----FKSSAQVQKFSR-----QSGGRGQYGDVHIE----- 534
 Qy 426 LGEFFQTKYDWDLLAARSIWAFGPDATG-----PNILVDDTLPSEVDKALLGSKVKSIVQ 480
 | | : | | | : | : | | : : : | : :
 Db 535 -----FTPNETGAGFEFENAIVGGVVPREYIPSVEAGLKDAMEN 573
 Qy 481 G 481
 |
 Db 574 G 574

Search completed: January 30, 2004, 11:35:27
 Job time : 56.2763 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 11:17:27 ; Search time 57.0837 Seconds
(without alignments)
2174.410 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSVKDSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description
<hr/>						

1	2491	98.6	850	4	Q8IXJ3	Q8ixj3	homo sapien
2	2409	95.3	974	13	Q8AVN2	Q8avn2	xenopus lae
3	1848	73.1	975	5	Q9VAX8	Q9vax8	drosophila
4	1764.5	69.8	974	5	Q23463	Q23463	caenorhabdi
5	1754.5	69.4	849	5	Q19070	Q19070	caenorhabdi
6	1528.5	60.5	987	10	Q9LNC5	Q9lnc5	arabidopsis
7	1335	52.8	983	3	O94316	O94316	schizosacch
8	1062	42.0	1235	5	Q8IJZ9	Q8ijz9	plasmodium
9	903.5	35.8	728	5	Q9BNW7	Q9bnw7	scolopendra
10	881.5	34.9	726	5	Q9BNX1	Q9bnx1	nipponopsal
11	872.5	34.5	660	5	Q9BNW9	Q9bnw9	polyxenus f
12	869	34.4	858	11	Q8C153	Q8c153	mus musculu
13	867	34.3	858	11	Q60423	Q60423	cricetulus
14	863	34.2	727	5	Q9BNW0	Q9bnw0	peripatus s
15	862.5	34.1	844	5	Q95P39	Q95p39	aedes aegyp
16	862.5	34.1	844	5	Q8T4R9	Q8t4r9	aedes aegyp
17	862	34.1	726	5	Q9BNX2	Q9bnx2	mastigoproc
18	861	34.1	702	5	Q9BNX6	Q9bnx6	endeis laev
19	856.5	33.9	844	5	Q9BME7	Q9bme7	aedes aegyp
20	855	33.8	858	11	Q8BMA8	Q8bma8	mus musculu
21	854.5	33.8	844	5	Q8T4S0	Q8t4s0	aedes aegyp
22	853.5	33.8	658	5	Q9BNX5	Q9bnx5	hutchinsoni
23	851.5	33.7	633	5	Q9BNY0	Q9bny0	artemia sal
24	851	33.7	727	5	Q9BNW8	Q9bnw8	scutigere
25	844	33.4	703	5	Q9BNW2	Q9bnw2	milnesium t
26	839	33.2	813	3	Q9P4S4	Q9p4s4	candida alb
27	838.5	33.2	658	5	Q9BNX4	Q9bnx4	limulus pol
28	838.5	33.2	726	5	Q9BNX7	Q9bnx7	eumesocampa
29	837.5	33.1	658	5	Q9BNW5	Q9bnw5	tomocerus s
30	836.5	33.1	633	5	Q9BNX3	Q9bnx3	machiloides
31	836	33.1	841	5	Q95UT8	Q95ut8	monosiga br
32	835	33.0	813	3	Q9P4S0	Q9p4s0	clavispora
33	830.5	32.9	637	5	Q9BNW6	Q9bnw6	speleonecte
34	829	32.8	813	3	Q9P4R9	Q9p4r9	candida par
35	828.5	32.8	633	5	Q9BNX8	Q9bnx8	semibalanus
36	826	32.7	813	3	Q9P4S1	Q9p4s1	candida tro
37	816.5	32.3	656	5	Q9BNW1	Q9bnw1	nereis vire
38	816	32.3	789	10	Q9M655	Q9m655	euglena gra
39	808	32.0	814	3	Q9P4S3	Q9p4s3	candida gla
40	804	31.8	838	3	Q9HFZ8	Q9hfhz8	cryptococcu
41	803	31.8	760	5	Q9NIH0	Q9nih0	stylonychia
42	802	31.7	753	5	Q9NDT2	Q9ndt2	plasmodium
43	802	31.7	832	5	Q8IKW5	Q8ikw5	plasmodium
44	801.5	31.7	726	5	Q9BNW4	Q9bnw4	tanystylum
45	793.5	31.4	665	10	Q8H145	Q8h145	arabidopsis

ALIGNMENTS

RESULT 1

Q8IXJ3

ID Q8IXJ3 PRELIMINARY; PRT; 850 AA.

AC Q8IXJ3;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Small nuclear ribonucleoprotein component.
GN SNRP116.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennerz V., Fatho M., Gentilini C., Lifke A., Woelfel C., Woelfel T.;
RT "Response of autologous T cells to a human melanoma is dominated by
RT individual mutated antigens.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.";
RL DNA Res. 1:27-35(1994).
DR EMBL; AJ505017; CAD43720.1; -.
KW Nucleocapsid; Ribonucleoprotein.
SQ SEQUENCE 850 AA; 95371 MW; D25DAEB8193D25D8 CRC64;

Query Match 98.6%; Score 2491; DB 4; Length 850;
Best Local Similarity 99.4%; Pred. No. 4.5e-192;
Matches 475; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy	4	YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIFYNP	63
Db	170	YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIFYNP	229
Qy	64	KTRKFTTKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKHEELKLN	123
Db	230	KTRKFTTKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKHEELKLN	289
Qy	124	IRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD	183
Db	290	IRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD	349
Qy	184	GPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW	243
		:	
Db	350	GPLMCHTTKMYSTDDGVQFHAFAFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW	409
Qy	244	ISVARYHIEVNRVPAGNWWVLIIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA	303
Db	410	ISVARYHIEVNRVPAGNWWVLIIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA	469
Qy	304	VEPVNPSELPMKLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI	363
Db	470	VEPVNPSELPMKLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI	529
Qy	364	DIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLEKGLAEDIENEVQITWNR	423
Db	530	DIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLEKGLAEDIENEVQITWNR	589

Qy 424 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 Db 590 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 647

RESULT 2

Q8AVN2

ID Q8AVN2 PRELIMINARY; PRT; 974 AA.
 AC Q8AVN2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to U5 snRNP-specific protein, 116 kD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC041724; AAH41724.1; -.
 SQ SEQUENCE 974 AA; 109777 MW; 35B07DCA5B23FE9D CRC64;

Query Match 95.3%; Score 2409; DB 13; Length 974;
 Best Local Similarity 94.6%; Pred. No. 2.3e-185;
 Matches 452; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

Qy 4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDYFNP 63
 Db 294 YSTDENLILSPLLGNVCFASSQYSICFTLGSAFAKIYADTYGDINYQEFAKRLWGDYFNP 353
 Qy 64 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLN 123
 Db 354 KTRKFTKKAPTSSSQRSFVEFVLEPLYKILAQVVGVDVDTLPQTLEELGIHLTKHEELKLN 413
 Qy 124 IRPLLRLVCKKFFGEFTGFVDMCVQHPSPKVGAKPKIEHTYTGGVSDLGAMSDCDPD 183
 Db 414 IRPLLRLVCNRRFFGEFTGFVDMCVQHPSPKAGARAKIEHTYTGGIDSELGEVMSECDPD 473
 Qy 184 GPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 243
 Db 474 GPLMCHTTKMYSTDDGVQFRAFGRLSGTIHAGQPVKVLGENYTLDEEDSQVCTIGRLW 533
 Qy 244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
 Db 534 VSVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 593
 Qy 304 VEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 363
 Db 594 VEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 653
 Qy 364 DIKVADPVVTFCEVTVETSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNR 423

Db 654 DIKVADPVVTFCEVTVETSSLKCF AETPNKKNKITMIAEPLEKGLAEDIENEVVQISWNR 713

Qy 424 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 |||||

Db 714 KKLGEFFQTKYDWDLLAARSIWAFGPDTTGPNILVDDTLPSEVDKALLSSVKDSIVQG 771

RESULT 3

Q9VAX8

ID Q9VAX8 PRELIMINARY; PRT; 975 AA.

AC Q9VAX8;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE CG4849 protein (LD28793p).

GN CG4849.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; *Drosophila*.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

```

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE003763; AAF56769.1; -.
DR EMBL; AY089551; AAL90289.1; -.
DR FlyBase; FBgn0039566; CG4849.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Protein biosynthesis.
SQ SEQUENCE 975 AA; 110649 MW; 68A9072B3300E9CB CRC64;

```

QY 303 AVEPVNPSELPMKLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSE 362
 |||:|
 Db 594 AVEPVNPSELPMKLDGLRKVNKSYPLLSTRVEESGEHVILGTGELYLDCVMHDLRKMYSE 653
 QY 363 IDIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVVQITWN 422
 |||:|
 Db 654 IDIKVADPVVAFCEVTVETSSSLKCAETPNKKNKITMISEPLEKGLAEDIENGTVGINWN 713
 QY 423 RKKLGEEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSSEVDKALLGSKVDSIVQG 481
 :|:|
 Db 714 KKRIGEEFFQVNYDWDLLAARSIWAFGPDSTGPNILVDDTLPSSEVDKNLLTAVKDSIVQG 772

RESULT 4

Q23463

ID Q23463 PRELIMINARY; PRT; 974 AA.
 AC Q23463;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 110.5 kDa protein.
 GN ZK328.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Favello A.;
 RT "The sequence of C. elegans cosmid ZK328.";
 RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; U50193; AAA91248.1; -.
 DR WormPep; ZK328.2; CE05066.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.

KW Hypothetical protein; GTP-binding; Protein biosynthesis.
 SQ SEQUENCE 974 AA; 110484 MW; 44C4A5FDA61F013C CRC64;

Query Match 69.8%; Score 1764.5; DB 5; Length 974;
Best Local Similarity 69.1%; Pred. No. 2.3e-133;
Matches 331; Conservative 66; Mismatches 81; Indels 1; Gaps 1;

Qy	4	YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGD-INYQEFAKRLWGDIIYFN	62
Db	295	FAEEDVPVLSPLNGNVIFSSGRYNVCFSLLSFSNIYAKQHGDSEFNSKEFARRLWGDIIYFE	354
Qy	63	PKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDDVDTSLPRTLDELGIHLTKEELKL	122
Db	355	KKTRKFVKKSPSHDAPRTFVQFILEPMYKIFSQVVGDDVDTCLPDVMAELGIRLSKEEQKM	414
Qy	123	NIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTTGGVSDSLGEAMSDCDP	182
Db	415	NVRPLIALICKRFFGDFSAFVDLVVQNIKSPLENAKTKEQTYLGPADSQLAQEMQKCNA	474
Qy	183	DGPLMCHTTKMFSTHDBGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRL	242
Db	475	EGPLMVHTTKNYPVDDATQFHVFGRVMSGTLEANTDVRVLGENYSIQDEEDCRRMTVGRL	534
Qy	243	WISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKI	302
Db	535	FVRVASYQIEVSRVPAGCWVLEIGIDQPIVKTATIAELGYEEDVYIFRPLKFNTRSCVKL	594
Qy	303	AVEPVNPSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYS	362
Db	595	AVEPINPSELPKMLDGLRKVNKSYPLLTTRVEESGEHVLLGTGEFYMDCMHDMRKVFSE	654
Qy	363	IDIKVADPVVTFCETVVETSSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVQITWN	422
Db	655	IDIKVADPVVTFNETVIETSTLKCFAETPNKKNKITMMAEPLKQLDEEDIENEVQIGWN	714
Qy	423	RKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSVDKALLGSVKDSIVQG	481
Db	715	RRRLGEFFQTKYNWDLLAARSIWAFGPDTTGPNILLDDTLPSVDKHLSTVRESLVQG	773

RESULT 5

Q19070

ID O19070 PRELIMINARY; PRT; 849 AA.

AC Q19070;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Eft-1 protein.

GN EFT-1.

OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=92153310; PubMed=1739435;

AC Q9LNC5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE F9P14.8 protein.
 GN F9P14.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S.X., Pham P.,
 RA Toriumi M., Yu G., Chin C., Chiou J., Choi E., Chung M., Gonzalez A.,
 RA Howng B., Liu A., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
 RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
 RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC F9P14 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AC025290; AAF80219.1; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 KW GTP-binding; Protein biosynthesis.
 SQ SEQUENCE 987 AA; 110441 MW; 868DD5F6597FBB79 CRC64;

Query Match 60.5%; Score 1528.5; DB 10; Length 987;
 Best Local Similarity 61.2%; Pred. No. 2.6e-114;
 Matches 289; Conservative 69; Mismatches 113; Indels 1; Gaps 1;

Qy 11 ILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFG-DINYQEFARKLWGDIFYNPKTRKFT 69
 Db 312 LIDPAAGNVCFASGTAGWSFTLQSFARKMYAKLHGVA MDVDKFA SRLWGDVYYHSDTRVFK 371
 Qy 70 KKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKLEELKLNIRPLLR 129
 Db 372 RSPPVGGGERAFVQFILEPLYKIYSQVIGEHHKKS VETTLAELGVTLNSAYKLNVRPLLR 431
 Qy 130 LVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSDLGEAMSDCDPDGPI MCH 189
 Db 432 LACSSVFGSASGFTDMLVKHIPS PREAAARKVDHSYTG TKDSPIYESMVECDPSGPLMVN 491
 Qy 190 TTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENY TLEDEEDSQICTVGRWLWISVARY 249

QY 5 STDENLILSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFARLWGDIFYNPK 64
 | | :| | | | | :| | | | | | :| | :| | | | | :|
 Db 304 SKDLKYRVSPELGNVCFASCDLGYCFTLSSFAKLYIDRHGGIDVDLFSKRLWGDIFYFDSK 363
 QY 65 TRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLN 124
 | | | | :| :| | | | | | | :| | :| | :| :| :| :|
 Db 364 TRKFAKQSLDGSVRSFVHFILEPLYKLHTLTISDEAEKLKKHLSSFQIYLKPKDYLLDP 423
 QY 125 RPLRLVCKKFFGFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLEAMSDC--DP 182
 :| | | :| :| | | :| | | :| | :| :| :| :| :| :|
 Db 424 KPLLQLICASFEGFPVGVNAVTRHIPS PRENAARKASQSYIGPINSSIGKAILEMSREE 483
 QY 183 DGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQICTVGRL 242
 | | :| | | :| :| | | :| | | | | :| | | | | :| :| :|
 Db 484 SAPLVMHVTKLYNTVDANNFYAFARVYSGQVKKGQKVKVLGENYSLEDEEDMVVAHIAEI 543
 QY 243 WISVARYHIEVNRVPAGNWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKI 302
 :| | | :| :| | | :| | | | | :| :| | | :| | | :|
 Db 544 CVPCARYRLHVDGAVAGMLVLLGGVDNSISKATIVSDNLKDDPYIFRPIAHMSSESVFKV 603
 QY 303 AVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSE 362
 | | | | | | :| | | | | | | | | | | | :| :| :| :| :| :|
 Db 604 AVEPHNPSELPKLLDGLRKTNKSYPSTTKVEESGEHTIFGTGEMYMDCLLYDLRTLYSE 663
 QY 363 IDIKVADPVVTFCEVTVETSSILKCFATPNKKNKITMIAEPLEKGLAEDIENEVQITWN 422
 :| :| :| :| | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 664 IEIRVSDPVARFCETAVDTSIKCFSDTPNKKNRITMVVEPLEKGISNDIENGKVNINWP 723
 QY 423 RKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 :| :| :| | | | | :| | | | | | | | | | :| :| :| :| :| :| :|
 Db 724 QKRISEFFQKNYDWDLLASRSIWAFGPDDRGTNLRDDTLSTDVDKNVLNSVKEYIKQG 782

RESULT 8

Q8IJZ9

ID Q8IJZ9 PRELIMINARY; PRT; 1235 AA.
 AC Q8IJZ9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE U5 small nuclear ribonuclear protein, putative.
 GN PF10_0041.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
 RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum.";
 RL Nature 419:498-511(2002).
 DR EMBL; AE014829; AAN35239.1; -.
 KW Nuclear protein.
 SQ SEQUENCE 1235 AA; 142943 MW; EEE0EE5592AFEB6E CRC64;

Query Match 42.0%; Score 1062; DB 5; Length 1235;
 Best Local Similarity 38.3%; Pred. No. 1.6e-76;
 Matches 235; Conservative 80; Mismatches 159; Indels 140; Gaps 7;

Qy 8 ENLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFG--DINYQEFKRLWGDIIYFNPKT 65
 :: :||| | | | | | : | | | | : : : | : | | | | : :
 Db 422 KDFLLSPKNNVLFSSSIYGVFFTLKSFSKIYCNIYNAYNIDIDEFSLYLWGDIIYDEEN 481

Qy 66 RKFTKKAPTSSSQRSFEFILEPLYKILAQVVGVDVTSLPRTLDELGIHLTKEELKLNIR 125
 | | : : | : | | | | | | | | | : | | | | : | :
 Db 482 FTFVKSPLYANQKRTFVEFILNPLYKIFGYVCSSEEKEFLIPFLQSFNISKKS DYLYNTK 541

Qy 126 PLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DLGEAMSDC---DP 182
 || : : | : | | | : : : | | | | : | : : : | | |
 Db 542 YLLKKINGMIFQDTTAFVDI IDNCPSPLDNAKNKTLQIYSGSLKTKISYDMMKCMKGDE 601

Qy 183 DGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEEDSQICTVGRL 242
 || : | : : | | | : | | | | : | | | | | : | | : |
 Db 602 TDNLMIYIIKNYHRPECGMIDLFGRVMCGTIKKGQSVRILGEGYTLNDDEDMITRVITHL 661

Qy 243 WISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPR----- 281
 || | | | : : | | : | | | | :
 Db 662 WIYEGRYRIEVDEIPAGNFC LIGGIDICINKTCTITNVKIKTNKNKEDHVNINWYDDLND 721

Qy 282 -----GNEEAQIFRPL--KFN----TTSVIKIAVEPVNPSE 311
 | | | : | | | | | | | | | : | | | | | |
 Db 722 DDNNIVKYKNDKDIYNLDKSALLNDNENAEIFYPLHKKFRYLNVCNSVFKVACEPINPSE 781

Qy 312 LPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPV 371
 || || | : | | : | | : | | | | | | | | | | | : | | | : | : | : | : | |
 Db 782 LPKMLEGLRKIDKVYPLSSTKVEESGEHIILGTGELYLDCILHDLRKLYGDLEIKVSDPV 841

Qy 372 VTFCETVETSSLKCAETPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRK-KLGEFF 430
 | | | : | : | | | | | | | | | : | | : | : : : | :
 Db 842 VQFNETIETSALNCFAETPNKKNKIYMIAPVQKELGDDIVQGLVHLNEDQNLNVNEYI 901

Qy 431 QT----- 432
 |
 Db 902 STMDRILNKNNESKRNIYDEKDDSHDDNDDGDNMENHMGNIENKKNIFKDKEQKDEHQIN 961

Qy 433 -----KYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVD 467
 | : | : | : | : | | | : | | | | : | :
 Db 962 DLDKEKRKSTLNYNIDPNVISLLKNKYNWDILSIRSLWAFGPENNSPNILVDDSLFQETN 1021

Qy 468 KALLGSVKDSIVQG 481
 | | | : | : | : |
 Db 1022 KENLYSIKDNIIQG 1035

09BNW7

Query Match 35.8%; Score 903.5; DB 5; Length 728;
Best Local Similarity 39.7%; Pred. No. 4.6e-64;
Matches 187; Conservative 99; Mismatches 176; Indels 9; Gaps 4;

Qy	12	LSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIYFNPKTRKFTK-	70
Db	201	VDPKSGSVGFGSGLHGWAFATLKQFSEIYAEKF-KIDVEKLMKRLWGENFYNP KSKKWAKS	259
Qy	71	KAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL	130
Db	260	NDETGEFKRSCFMIILDPIYKVFDAIMNYKTDEIPKLEKLNIIKGEDKDKDGKSLKLV	319
Qy	131	VCKKFFGEFTGFVDMCVQHIPS PKVGA PKIEHTYTGGVSDSLGEAMSDCDPDG PLMCHT	190
Db	320	VMRQWLPAGEALLQMI AIHLSPVKAQKYRMEMLYEGPHDDEAAVAVKNCDP SGPLMMYV	379
Qy	191	TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEEDSQICTVGRWLWISVARYH	250
Db	380	SKMVPTSDKGRFYAFGRVFSGVVSTGQKVRIMGPNYTPGKKEDLYEKAIQRTILMMGRYT	439
Qy	251	IEVNRVPAGNWWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS	310
Db	440	EAIEDVPCGNICGLVGVDQFLVKTGTITT---FKDAHNMRVMKFSVSPVVRVAVEPKNPS	496

Qy	311	ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLD CVMHDLRKMYS EIDIKVADP	370
		: ::: ::: : : :: ::: : : :: ::	
Db	497	DLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPIKVS DP	556
Qy	371	VVTFCE TVVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIEN EVVQITWNRKKLG EFF	430
		:: : :: :: ::: : :: : :	
Db	557	VVS YRET VSE ESEIMCLAKSPNKHNR LFMKAQPMPEGLPEDIDKGDVTARDDFKARARYL	616
Qy	431	QTKYDWDLLAARS IWA FGP DATGPNILVDDTLPSEVDKALLG SVKDSIVQG	481
		::: : ::: : : ::	
Db	617	SDKYDYDVTEARKIWC FGP DGTGPNILIDCTKG VQ---YLNEIKDSV VAG	663

RESULT 10

Q9BNX1

```

ID      Q9BNX1          PRELIMINARY;          PRT;      726 AA.
AC      Q9BNX1;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Elongation factor-2 (Fragment).
OS      Nipponopsalis abei.
OC      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Opiliones;
OC      Palpatores; Troguloidea; Nipponopsalididae; Nipponopsalis.
OX      NCBI_TaxID=109753;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21317060; PubMed=11421654;
RA      Regier J.C., Shultz J.W.;
RT      "Elongation factor-2: a useful gene for arthropod phylogenetics.";
RL      Mol. Phylogenet. Evol. 20:136-148(2001).
DR      EMBL; AF240824; AAK12349.1; -.
DR      HSSP; P13551; 1FNM.
DR      InterPro; IPR005517; EFG_IV.
DR      InterPro; IPR004161; EFTU_D2.
DR      InterPro; IPR000795; EF_GTPbind.
DR      InterPro; IPR005225; Small_GTP.
DR      Pfam; PF03764; EFG_IV; 1.
DR      Pfam; PF00009; GTP_EFTU; 1.
DR      Pfam; PF03144; GTP_EFTU_D2; 1.
DR      PRINTS; PR00315; ELONGATNFCT.
DR      TIGRFAMs; TIGR00231; small_GTP; 1.
DR      PROSITE; PS00301; EFATOR_GTP; 1.
KW      GTP-binding; Protein biosynthesis.
FT      NON_TER      1      1
FT      NON_TER      726      726
SQ      SEQUENCE      726 AA; 81041 MW; 8EADFAFDD2C3DED CRC64;

```

Query Match 34.9%; Score 881.5; DB 5; Length 726;
Best Local Similarity 37.9%; Pred. No. 2.7e-62;
Matches 178; Conservative 108; Mismatches 175; Indels 9; Gaps 4;

QY 12 LSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWGDIFYNPKTRKFTKK 71
: | |: | | : | || |: ||: | |: : : |||: : ||::: ||
Db 201 VDPSKGSVGFSGSLHSWAFTLKQFSEIYAEKF-KIDVEKLMNRLWGENFYNPQSKKWSKK 259

Qy	72	APTSSSQRSFEVFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLRLV	131
Db	260	M-DEGFKRAFCMFVLDPIYKIFKAIMGYQKEETAKLLEKLNIIILKGDDKEKDGNLLKVV	318
Qy	132	CKKFFGGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHTT	19
Db	319	MRNWLPA GDALLQMI AIHLPSPVTAQRYRIDLLYEGPQDDEAAVAMKSCDPDGPLMMYIS	378
Qy	192	KMFSTHDGVQFHFPGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQICTVGRLWISVARYHI	251
Db	379	KMVPTSDKGRFYAFGRVFSGIVSSGQKVRIMGPNYLP GK KDDLAEKAIQRTVLM GRAVE	438
Qy	252	EVNRVPAGNWWLIEGVDQPIVKTTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE	311
Db	439	PIENVPSGNICGLVGVDQFLVKTGTIST---FKDAHNMRVMKFSVSPVVRVAVEPMNPSD	495
Qy	312	LPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEIDIKVADPV	371
Db	496	LPKLVEGLKRLAKSDPMVQCII EESGEHIVAGAGELHLEICLKDLEEDHAGIPLKKTDPV	555
Qy	372	VTFCETVVETSSLKCAETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNRKKLGEFFQ	431
Db	556	VTYRETVAEESA IMCLSKSPNKHNRLYMKATPMQDGLPEDIDSGAVNPKDDFKARGYLS	615
Qy	432	TKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKALLGSVKDSIVQG	481
Db	616	DKYEWDAT EARKIWC FGP EGTGPNLLVDVT KGVO----YLNEIKDSV VAG	661

FT NON_TER 1 1
 FT NON_TER 660 660
 SQ SEQUENCE 660 AA; 73957 MW; 303B7891D8656E8C CRC64;

Query Match 34.5%; Score 872.5; DB 5; Length 660;
 Best Local Similarity 38.2%; Pred. No. 1.2e-61;
 Matches 179; Conservative 103; Mismatches 177; Indels 9; Gaps 4;

Qy 12 LSPLLGNVCFSSSQYSICFTLGSAFIYADTFGDINYQEFARLWGDYFNPTRKFTKK 71
 : | | : | | : | | | : : : | | : : : | | : : : | |
 Db 201 VDPSKGSVGFSGSLHGWAFTLKQFSELYAEKFG-IDVDKLMRRLWGENFYNPSSKKWAKS 259
 Qy 72 APTSSS-QRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL 130
 : : | | | : : : | | : : : | | : : : | | : : : | |
 Db 260 SNEGPDFKRSFCMFVLDPIYKVFDAMNYKTEEIPKLEKLNIVLKGEDKEKDGLLLKT 319
 Qy 131 VCKKFFGEFTGFVDMCVQHIPS PKVGA PKIEHTYTG GVDSDLGEAMSDCDPDG PLMCHT 190
 | : : : : | : | | | : : | | : : : | | : : : | | :
 Db 320 VMRQWLPAGEALLQMITIHLPSPVTAQKYRMELLYEGPHDDEAALGIKNCDP NGPLMMYV 379
 Qy 191 TKMFSTHDGVQFHFPGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRWLWISVARYH 250
 : | | | : : | | | : | | : : : | | : : : | | : : : | |
 Db 380 SKMVPTTDKGRFYAFGRVFSGIVSTGQKVRIMGPNYTPGKKEDLYEKAIQRTILMMGRYT 439
 Qy 251 IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
 : | | | : | | | : | | : : : | | : : : | | : : : | |
 Db 440 EAIEEVPCGNICGLVGVDQFLVKTGTIST---FKDAHNLRVMKFSVSPVVRVAVEAKNPS 496
 Qy 311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADP 370
 : | | : : : | | : : : | | : : : | | : : : | | : : : | |
 Db 497 DLPKLVEGLKRLAKSDPMVQCII EESGEHIIAGAGELHLEICLDLEEDHACIPIKVSDP 556
 Qy 371 VVTFCE TVVETSS LKCF AETPNKKNKITMIAEPLEKGLAEDIENEVVQITWNRKKLGEFF 430
 | | : | | | : | | : | | : | | : | | : | | : | | : | |
 Db 557 VVSRET VTEESDIMCLAKSPNKHNRLYMKAVPMPEGLPEDIDKGEVTARDDFKSRGRL 616
 Qy 431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSKVSDSI 478
 | | : : | | | | | | : | | : | | : | | : | | : | |
 Db 617 AEKYEYDITEARKIWCFGPDGTGPNILIDCTKG VQ---YLNEIKDSV 660

RESULT 12

Q8C153

ID Q8C153 PRELIMINARY; PRT; 858 AA.
 AC Q8C153;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Elongation factor 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;

GN EF2.
 OS *Cricetulus griseus* (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC *Cricetulus*.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96032765; PubMed=7559470;
 RA Foley B.T., Moehring J.M., Moehring T.J.;
 RT "Mutations in the elongation factor 2 gene which confer resistance to
 RT diphtheria toxin and *Pseudomonas* exotoxin A. Genetic and biochemical
 RT analyses.";
 RL J. Biol. Chem. 270:23218-23225(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=86259716; PubMed=3014523;
 RA Kohno K., Uchida T., Ohkubo H., Nakanishi S., Nakanishi T., Fukui T.,
 RA Ohtsuka E., Ikehara M., Okada Y.;
 RT "Amino acid sequence of mammalian elongation factor 2 deduced from the
 RT cDNA sequence: homology with GTP-binding proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4978-4982(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=88198187; PubMed=2834376;
 RA Nakanishi T., Kohno K., Ishiura M., Ohashi H., Uchida T.;
 RT "Complete nucleotide sequence and characterization of the 5'-flanking
 RT region of mammalian elongation factor 2 gene.";
 RL J. Biol. Chem. 263:6384-6391(1988).
 DR EMBL; U17362; AAB60497.1; -.
 DR HSSP; P13551; 1FNM.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW GTP-binding; Protein biosynthesis.
 FT CONFLICT 441 441 D -> E (IN REF. 1,3,2,2 AND 3).
 FT VARIANT 584 584 S -> G.
 FT VARIANT 714 714 I -> N.
 FT VARIANT 717 717 G -> R.
 FT VARIANT 719 719 G -> D.
 SQ SEQUENCE 858 AA; 95310 MW; 3F103E20D69140D8 CRC64;

Query Match 34.3%; Score 867; DB 11; Length 858;
 Best Local Similarity 37.3%; Pred. No. 5.1e-61;
 Matches 182; Conservative 94; Mismatches 190; Indels 22; Gaps 5;

DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW GTP-binding; Protein biosynthesis.
 FT NON_TER 1 1
 FT NON_TER 727 727
 SQ SEQUENCE 727 AA; 81394 MW; A7E368E3CFDE8FDB CRC64;

Query Match 34.2%; Score 863; DB 5; Length 727;
 Best Local Similarity 37.6%; Pred. No. 8.4e-61;
 Matches 182; Conservative 101; Mismatches 189; Indels 12; Gaps 4;

QY 2 AYYSTDE----NLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFARLWG 57
 | | : : : | | | | : | | | | : | : | : | | |
 Db 187 ATYSDESGPMGDIKVDPSKGNVGLGSLHGWAFTLKQFAEIYSEKF-KIDVDKLMKRLWG 245
 QY 58 DIYFNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDDVDTSLPRTLDELGIHLTK 117
 : : | | | | : : | : : : | : | : | : | : : : | : | | |
 Db 246 ENFYNPARKWSKKCESEDYKRAFCMFVLDPIYKIFDAIMNYKKDETAKLLEKLNIVLKG 305
 QY 118 EELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDLGEAM 177
 | : : : | : : : : : : | : | | : : | | | : : : |
 Db 306 EDKDKDGKALLKIVMRTWLPAGEALLQMIALHLPSPVTAQRYRTELLYEGPQDDEVAISM 365
 QY 178 SDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDDSQIC 237
 : | : | | : : : | | : : | | | | : : | | : : | | : | |
 Db 366 KECNPQGGLIMYISKMVPTSDKGRFYAFGRVFSGLVSTGQKVRIMGPNYVPGKKEDLYEK 425
 QY 238 TVGRLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTT 297
 : | : : | | : | | | : | | | | | | : : | | : | : :
 Db 426 AIQRTILMMGRYVEAIEDVPCGNICGLVGVDQFIVKTGTITT---FKDAHNLRVMKFSVS 482
 QY 298 SVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLR 357
 | : : | | | | | : : : : : : : | : : | | | | : : | |
 Db 483 PVVRVAVEAKNPSDLPKLVEGLKRLSKSDPMVQCIIIESGEHIIAGAGELHLEICLKDLE 542
 QY 358 KMYSEIDIKVADPVVTFCETVVETSSLKCFETPNKKNKITMIAEPLKGLAEDIENEV 417
 : : : | | | : | | : | | : : | | : | | : | | : |
 Db 543 EDHAGIPIKVS DPVVS YRETVSDESDTMCLSKSPNKHNRLEFMKAVPMPDGLPEDIDKGEV 602
 QY 418 QITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDS 477
 | : : | | : | | | | : : | | | | : : | : | : |
 Db 603 SPKGEFKARARYLGEKYDYDVSEARKIWCFGPEGSGPNILVDCTKG VQ----YLNEIKDS 658
 QY 478 IVQG 481
 : | |
 Db 659 VVAG 662

RESULT 15

Q95P39

ID Q95P39 PRELIMINARY; PRT; 844 AA.

Db 503 LPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPLKKSDPV 562
 Qy 372 VTFCETVVETSSLKCFETPNKKNKITMIAEPLEKGLAEDIENEVQITWNRKKLGEFFQ 431
 |:: ||| : | | ::||| |:: | | |: |||||:| | : | :
 Db 563 VSYRETVSDESDQMCLSKSPNKHNRLEFMKAVPMPDGLAEDIDNGDVNSRDDFKVRARYLA 622
 Qy 432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 |||:|: || || |||| ||||:| | : | :|||:| |
 Db 623 EKYDYDVTEARKIWCFGPDGTGPNIVVDCTKGVQ-----YLNEIKDSVVAG 668

Search completed: January 30, 2004, 11:26:21
 Job time : 59.0837 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 11:15:32 ; Search time 18.716 Seconds
(without alignments)
1208.586 Million cell updates/sec

Title: US-09-989-481-3
Perfect score: 2527
Sequence: 1 RAYYSTDENLILSPLLGNVC.....LPSEVDKALLGSVKDSIVQG 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2491	98.6	972	1	U5S1_HUMAN	Q15029 homo sapien
2	2487	98.4	971	1	U5S1_MOUSE	O08810 mus musculu
3	870	34.4	857	1	EF2_HUMAN	P13639 homo sapien
4	869	34.4	857	1	EF2_MOUSE	P58252 mus musculu
5	868	34.3	857	1	EF2_RAT	P05197 rattus norv
6	867	34.3	857	1	EF2_CRIGR	P09445 cricetulus
7	865	34.2	857	1	EF2_CHICK	Q90705 gallus gall
8	863	34.2	857	1	EF2_MESAU	P05086 mesocricetu
9	854	33.8	832	1	EF2_CRYPV	Q23716 cryptospori
10	844.5	33.4	843	1	EF2_DROME	P13060 drosophila
11	839	33.2	842	1	EF2_CANAL	O13430 candida alb
12	837.5	33.1	851	1	EF2_CAEEL	P29691 caenorhabdi
13	823.5	32.6	840	1	EF2_ENTHI	Q06193 entamoeba h
14	816.5	32.3	843	1	EF2_BETVU	O23755 beta vulgar
15	807.5	32.0	842	1	EF2_YEAST	P32324 saccharomyc
16	801.5	31.7	842	1	EF2_SCHPO	O14460 schizosacch
17	788.5	31.2	844	1	EF2_NEUCR	Q96x45 neurospora

18	779	30.8	845	1	EF2_CHLKE	P28996	chlorella k
19	751.5	29.7	867	1	EF2_BLAHO	Q17152	blastocysti
20	710.5	28.1	830	1	EF2_DICDI	P15112	dictyosteli
21	668.5	26.5	1008	1	SN14_YEAST	P36048	saccharomyc
22	404.5	16.0	736	1	EF2_AERPE	Q9yc19	aeropyrum p
23	396.5	15.7	732	1	EF2_PYRHO	O59521	pyrococcus
24	390.5	15.5	728	1	EF2_ARCFU	O28385	archaeoglob
25	390.5	15.5	1110	1	YNQ3_YEAST	P53893	saccharomyc
26	385.5	15.3	732	1	EF2_PYRAB	Q9v1z8	pyrococcus
27	385.5	15.3	732	1	EF2_PYRFU	P29050	pyrococcus
28	371.5	14.7	740	1	EF2_PYRAE	Q8zzc1	pyrobaculum
29	370	14.6	736	1	EF2_SULTO	Q975h5	sulfolobus
30	366.5	14.5	735	1	EF2_SULSO	P30925	sulfolobus
31	335	13.3	736	1	EF2_SULAC	P23112	sulfolobus
32	331	13.1	726	1	EF2_METJA	Q58448	methanococc
33	326.5	12.9	730	1	EF2_METTH	O27131	methanobact
34	326.5	12.9	732	1	EF2_THEAC	P26752	thermoplasm
35	322.5	12.8	734	1	EF2_DESMO	P33159	desulfuroco
36	318.5	12.6	727	1	EF2_METVA	P09604	methanococc
37	305.5	12.1	730	1	EF2_METBU	O93632	methanococc
38	303	12.0	728	1	EF2_HALHA	P14823	halobacteri
39	296	11.7	730	1	EF2_METMA	Q8pur7	methanosarc
40	295	11.7	730	1	EF2_METTE	O93640	methanosarc
41	290	11.5	730	1	EF2_METAC	Q8trc3	methanosarc
42	285	11.3	730	1	EF2_METMT	O93637	methanococc
43	225.5	8.9	698	1	EFG_DEIRA	Q9rxk5	deinococcus
44	219	8.7	695	1	EFG_LISMO	Q8y421	listeria mo
45	216	8.5	695	1	EFG_LISIN	Q927i5	listeria in

ALIGNMENTS

RESULT 1

U5S1_HUMAN

ID U5S1_HUMAN STANDARD; PRT; 972 AA.
AC Q15029; Q9BUR0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 116 kDa U5 small nuclear ribonucleoprotein component (U5 snRNP-specific protein, 116 kDa) (U5-116 kDa).
GN SNRP116 OR KIAA0031.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.";

RL DNA Res. 1:27-35(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=97377047; PubMed=9233818;
 RA Fabrizio P., Lagerbauer B., Lauber J., Lane W.S., Luhrmann R.;
 RT "An evolutionarily conserved U5 snRNP-specific protein is a GTP-
 RT binding factor closely related to the ribosomal translocase EF-2."
 RL EMBO J. 16:4092-4106(1997).
 CC -!- FUNCTION: COMPONENT OF THE U5 SNRNP COMPLEX REQUIRED FOR PRE-MRNA
 CC SPLICING. BINDS GTP.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D21163; BAA04699.1; -.
 DR EMBL; BC002360; AAH02360.1; -.
 DR SWISS-2DPAGE; Q15029; HUMAN.
 DR GK; Q15029; -.
 DR MIM; 603892; -.
 DR GO; GO:0005681; C:spliceosome complex; TAS.
 DR GO; GO:0003924; F:GTPase activity; TAS.
 DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.
 DR GO; GO:0006371; P:mRNA splicing; TAS.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.

DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; FALSE_NEG.
 KW GTP-binding; mRNA splicing; Nuclear protein; Polymorphism.
 FT NP_BIND 136 143 GTP (POTENTIAL).
 FT NP_BIND 204 208 GTP (POTENTIAL).
 FT NP_BIND 258 261 GTP (POTENTIAL).
 FT VARIANT 773 773 G -> V (IN dbSNP:1056505).
 FT /FTid=VAR_014931.
 FT CONFLICT 321 321 G -> V (IN REF. 2).
 SQ SEQUENCE 972 AA; 109435 MW; 862BD6CA7993F118 CRC64;

Query Match 98.6%; Score 2491; DB 1; Length 972;
 Best Local Similarity 99.4%; Pred. No. 3.5e-175;
 Matches 475; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDYFNP 63
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 292 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDYFNP 351

 Qy 64 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKEELKLN 123
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 352 KTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKEELKLN 411

 Qy 124 IRPLLRLLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 183
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 412 IRPLLRLLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPD 471

 Qy 184 GPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 243
 |||||||||:| ||||| ||||||||||||||||||||||||||||||||||||
 Db 472 GPLMCHTTKMYSTDDGVQFHAFAFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 531

 Qy 244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 532 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 591

 Qy 304 VEPVNPSELPMKLDGLRKVNKSYPSTTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI 363
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 592 VEPVNPSELPMKLDGLRKVNKSYPSTTTKVEESGEHVILGTGELYLDCVMHDLRKMYSI 651

 Qy 364 DIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNR 423
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 652 DIKVADPVVTFCEVTVETSSSLKCAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNR 711

 Qy 424 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 481
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 712 KKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDSIVQG 769

RESULT 2

U5S1_MOUSE

ID U5S1_MOUSE STANDARD; PRT; 971 AA.
AC O08810;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 116 kDa U5 small nuclear ribonucleoprotein component (U5 snRNP-specific protein, 116 kDa) (U5-116 kDa).
GN SNRP116.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97377047; PubMed=9233818;
RA Fabrizio P., Lagerbauer B., Lauber J., Lane W.S., Luehrmann R.;
RT "An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2.";
RL EMBO J. 16:4092-4106(1997).
RN [2]
RP SEQUENCE OF 401-971 FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: COMPONENT OF THE U5 SNRNP COMPLEX REQUIRED FOR PRE-MRNA
CC SPLICING. BINDS GTP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U97079; AAC53299.1; -.
 DR EMBL; BC012636; AAH12636.1; -.
 DR MGD; MGI:1336880; Snrpl16.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; FALSE_NEG.
 KW GTP-binding; mRNA splicing; Nuclear protein.
 FT NP_BIND 135 142 GTP (POTENTIAL).
 FT NP_BIND 203 207 GTP (POTENTIAL).
 FT NP_BIND 257 260 GTP (POTENTIAL).
 FT CONFLICT 401 404 HLTK -> PRVR (IN REF. 2).
 SQ SEQUENCE 971 AA; 109360 MW; 0ECF1661DEA3A7FC CRC64;

Query Match 98.4%; Score 2487; DB 1; Length 971;
 Best Local Similarity 99.2%; Pred. No. 6.9e-175;
 Matches 474; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIYFNP 63
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 291 YSTDENLILSPLLGNVCFSSSQYSICFTLGSAFAKIYADTFGDINYQEFAKRLWGDIYFNP 350

 QY 64 KTRKFTHKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKHEELKLN 123
 |||||:||||||||||||||||||||||||||||||||||||||||||
 Db 351 KTRKFTHKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDSLPRTLDELGIHLTKHEELKLN 410

 QY 124 IRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTTGGVSDSLGEAMSDCDPD 183
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 411 IRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTTGGVSDSLGEAMSDCDPD 470

 QY 184 GPLMCHTTKMFSTHGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 243
 |||||:||||||| |||||||| ||||||||||||||||||||||||
 Db 471 GPLMCHTTKMYSTDDGVQFHA FGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLW 530

 QY 244 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 303
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 531 ISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIA 590

 QY 304 VEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 363
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 591 VEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVMHDLRKMYSEI 650

 QY 364 DIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNR 423
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 651 DIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAE PLEKGLAEDIENEVVQITWNR 710

 QY 424 KKLGEFFQTKYDWDLLAARSIWAFGP DATGPNILVDDTL PSEVDKALLG SVKDSIVQG 481
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 711 KKLGEFFQTKYDWDLLAARSIWAFGP DATGPNILVDDTL PSEVDKALLG SVKDSIVQG 768

RESULT 3

EF2_HUMAN

ID EF2_HUMAN STANDARD; PRT; 857 AA.
AC P13639;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor 2 (EF-2).
GN EEF2 OR EF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=90121741; PubMed=2610926;
RA Rapp G., Klaudiny J., Hagendorff G., Luck M.R., Heinz K.;
RT "Complete sequence of the coding region of human elongation factor 2
RT (EF-2) by enzymatic amplification of cDNA from human ovarian
RT granulosa cells.";
RL Biol. Chem. Hoppe-Seyler 370:1071-1075(1989).
RN [2]
RP SEQUENCE OF 500-857 FROM N.A.
RX MEDLINE=88293714; PubMed=2840927;
RA Rapp G., Mucha J., Einspanier R., Luck M., Scheit K.H.;
RT "Cloning and sequence analysis of a cDNA from human ovarian granulosa
RT cells encoding the C-terminal part of human elongation factor 2.";
RL Biol. Chem. Hoppe-Seyler 369:247-250(1988).
RN [3]
RP CLEAVAGE OF INITIATOR METHIONINE.
RA Bienvenut W.V.;
RL Unpublished observations (AUG-2001).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19997; AAA50388.1; -.
DR EMBL; X51466; CAA35829.1; -.
DR EMBL; Z11692; CAA77750.1; -.

DR PIR; S18294; EFHU2.
 DR HSSP; P13551; 1FNM.
 DR PMMA-2DPAGE; P13639; -.
 DR Genew; HGNC:3214; EEF2.
 DR MIM; 130610; -.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT INIT MET 0 0
 FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 103 107 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 714 714 DIPHTHAMIDE.
 SQ SEQUENCE 857 AA; 95207 MW; 66421C8133EAC22E CRC64;

Query Match 34.4%; Score 870; DB 1; Length 857;
 Best Local Similarity 37.3%; Pred. No. 3.9e-56;
 Matches 182; Conservative 95; Mismatches 189; Indels 22; Gaps 5;

QY 9 NLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTF-----GDI-----NYQEFKRLW 56
 |::: |:|| | | : ||| |::| | : :: |::|
 Db 201 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 260
 QY 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
 || |::| |::| | : : |::| : |::|::| : : : : | |
 Db 261 GDRYFDPANGKFSKSATSPEGKKLPRTFCQLILDPIFKVFDAMNFKKEETAKLIEKLDI 320
 QY 114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
 | | : : |::| : |::| : | : |::| | : | | | :
 Db 321 KLDSEDKDKEGKPLLKAVMRRLPAGDALLQMITIHLPSPVTAQKYRCCELLYEGPPDDEA 380
 QY 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEED 233
 : ||| ||| : : | | : |::| || : | |::| || : | |
 Db 381 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAEGRVFSGLVSTGLKVRIMGPNYTPGKKED 440
 QY 234 SQICTVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
 : : | : : | : | | | : | | | : | | | | : | : |
 Db 441 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 497
 QY 294 FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDCVM 353
 | : : |::| |::|::|::|::|::| : : : : |::|::| : | : : :
 Db 498 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICL 557
 QY 354 HDLRKMYSEIDIKVADPVVTFCETVVETSSSLKCAETPNKKNKITMIAEPLEKGLAEDIE 413
 | | : : | | : |::|::| | | : | : : | | : : | | | | : :

Db 558 KDLEEDHACIPIKKS DPVVS YRETVSEES NVLCLSKSPNKHNR LYM KARPF PDGLAEDID 617

QY 414 NEVVQITWNRKKLGFEFFQTKYDWDL LAARSIWAFGPDATGPNILVDDTL PSEVDKALLGS 473

Db 618 KGEVSARQELKQRARYLA EKYEWDAEARKIWC FGPDGTGPNILTDITKG VQ----YLNE 673

QY 474 VKDSIVQG 481

Db 674 IKDSV VAG 681

RESULT 4

EF2_MOUSE

ID EF2_MOUSE STANDARD; PRT; 857 AA.

AC P58252;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Elongation factor 2 (EF-2).

GN EEF2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION

CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE

CC RIBOSOME.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2

CC (By similarity).

CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-

CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-

CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A (By

CC similarity).

CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; BC007152; AAH07152.1; -.
DR MGD; MGI:95288; Eef2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACTOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT NP_BIND 25 32 GTP (BY SIMILARITY).
FT NP_BIND 103 107 GTP (BY SIMILARITY).
FT NP_BIND 157 160 GTP (BY SIMILARITY).
FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 714 714 DIPHTHAMIDE (BY SIMILARITY).
SQ SEQUENCE 857 AA; 95182 MW; 1B96D46A8B9B7A44 CRC64;

Query Match 34.4%; Score 869; DB 1; Length 857;
Best Local Similarity 37.7%; Pred. No. 4.6e-56;
Matches 184; Conservative 93; Mismatches 189; Indels 22; Gaps 5;

Qy 9 NLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTF-----GDINYQEFA-----KRLW 56
|::: |:| | | : ||| |::| | : : | | :|
Db 201 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLSAAERAKKVEDMMKKLW 260

Qy 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
|| |:| |:| | : : |:| : |:|::| : : :::| |
Db 261 GDRYFDPANGKFSKANS PDGKKLPRTFCQLILDPIFKVFDAIMNFRKEETAKLIEKLDI 320

Qy 114 HLTKEELKLNIRPLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
| |: :||: |:: : | :||| | : | | | :
Db 321 KLDSEDKDKEGKPLLKAVMRRWL PAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 380

Qy 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEED 233
: ||| ||| : :| | | :| ||| | : | :::| ||| :|
Db 381 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKED 440

Qy 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
: : | : : | : || | | : ||| :||| ||| | | :|

RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE OF 701-715.
 RC TISSUE=Liver;
 RX MEDLINE=74301260; PubMed=4368673;
 RA Robinson E.A., Henriksen O., Maxwell E.S.;
 RT "Elongation factor 2. Amino acid sequence at the site of adenosine
 RT diphosphate ribosylation."
 RL J. Biol. Chem. 249:5088-5093(1974).
 RN [6]
 RP SEQUENCE OF 2-18 AND 50-59, AND PHOSPHORYLATION.
 RC TISSUE=Pancreas;
 RX MEDLINE=88087003; PubMed=3693353;
 RA Nairn A.C., Palfrey H.C.;
 RT "Identification of the major Mr 100,000 substrate for calmodulin-
 RT dependent protein kinase III in mammalian cells as elongation
 RT factor-2."
 RL J. Biol. Chem. 262:17299-17303(1987).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
 CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
 CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
 CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K03502; AAA41106.1; -.
 DR EMBL; Y07504; CAA68805.1; -.
 DR EMBL; U75403; AAB19107.1; -.
 DR EMBL; AF000576; AAD05363.1; -.
 DR PIR; S04007; EFRT2.
 DR HSSP; P13551; 1FNM.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT INIT_MET 0 0 BY SIMILARITY.

FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 103 107 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 714 714 DIPHTHAMIDE.
 SQ SEQUENCE 857 AA; 95152 MW; C808AD8ABD527E85 CRC64;

Query Match 34.3%; Score 868; DB 1; Length 857;

Best Local Similarity 37.7%; Pred. No. 5.4e-56;

Matches 184; Conservative 92; Mismatches 190; Indels 22; Gaps 5;

QY 9 NLILSPLLGNVCFSSSQYSICFTLGSEFAKIYADTF-----GDINYQEFA-----KRLW 56
 |::: |:| | | : | | |::| | : | | |::|
 Db 201 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGAAERAKKVEDMMKKLW 260
 QY 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGI 113
 || |:| |:| | : : |:| : |:|::| : : :::| |
 Db 261 GDRYFDPANGKFSKANSNPDGKKLPRTFCQLILDPIFKVFDAMNFRKEETAKLIEKLDI 320
 QY 114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
 | |: :||: |:: : | :|| | : | | | :
 Db 321 KLDSEDKDKEGKPLLKAVMRRLPAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 380
 QY 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
 : ||| ||| : :| | | :| ||| || : | :::| ||| :||
 Db 381 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKED 440
 QY 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
 : : | : : || : || || | : ||| :|| || | | | :|
 Db 441 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 497
 QY 294 FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLD CVM 353
 |: : |:::| | |::| |::| |::| | | : : |||||:| | |||:| :
 Db 498 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCTIEESGEHIIAGAGELHLEICL 557
 QY 354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSSLKCFAETPNKKNKITMIAE PLEKGLAEDIE 413
 || : : | | | :|||: || | |:: | :::|| |:: | | | |||||:
 Db 558 KDLEEDHACIPIKKS DPVVS YRETVSEESNVLCLSKSPNKHNRLYMKARFPDGLAEDID 617
 QY 414 NEVVQITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
 | | : |::|: || || ||| ||||| | | : |
 Db 618 KGEVSARQELKARARYLAEKYEWDAEARKIWCFGPDGTGPNILTDITKGVQ----YLNE 673
 QY 474 VKDSIVQG 481
 :|||:| |
 Db 674 IKDSVVAG 681

RESULT 6

EF2_CRIGR

ID EF2_CRIGR STANDARD; PRT; 857 AA.

AC P09445;

DT 01-MAR-1989 (Rel. 10, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Elongation factor 2 (EF-2).

GN EEF2.
 OS *Cricetulus griseus* (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC *Cricetulus*.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88198187; PubMed=2834376;
 RA Nakanishi T., Kohno K., Ishiura M., Ohashi H., Uchida T.;
 RT "Complete nucleotide sequence and characterization of the 5'-flanking
 RT region of mammalian elongation factor 2 gene.";
 RL J. Biol. Chem. 263:6384-6391(1988).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
 CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
 CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
 CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; J03200; AAA50386.1; -.
 DR HSSP; P13551; 1FNM.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 103 107 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 714 714 DIPHTHAMIDE.
 SQ SEQUENCE 857 AA; 95264 MW; DCE0A55CFBCB8886 CRC64;

Query Match

34.3%; Score 867; DB 1; Length 857;

Best Local Similarity 37.3%; Pred. No. 6.4e-56;
Matches 182; Conservative 94; Mismatches 190; Indels 22; Gaps 5;

Qy	9	NLILSPLLGNVCFSSSQYSICFTTGLGSFAKIYADTF-----GDI-----NYQEFAKRLW	56
Db	201	NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW	260
Qy	57	GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGI	113
Db	261	GDRYFDPANGKFSKANSAPDGKKLPRTFCQLILDPIFKVFDPIMNFRKEETAKLIEKLDI	320
Qy	114	HLTKEELKLNIRPLLRVLCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDSDL	173
Db	321	KLDSKDKKEGKPLLKAVMRRWLPA GDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA	380
Qy	174	GEAMSDCDPDGPILMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLEDEED	233
Db	381	AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKED	440
Qy	234	SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK	293
Db	441	LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTTGTTT---FEHAHNMRVMK	497
Qy	294	FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYP SLTTKVEESGEHVILGTGELYLDCVM	353
Db	498	FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCII EESGEHIIAGAGELHLEICL	557
Qy	354	HDLRKMYSEIDIKVADPVVTF CETVVETSSLKCF AETPNKKNKITMIAEPLEKGLAEDIE	413
Db	558	KDLEEDHACIPIKKSDPVVSYRET VSEESNVLCLSKSPNKHNRLYMKARFPDPGLAEDID	617
Qy	414	NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS	473
Db	618	KGEVSARQELKARARYLAEKYEWDVAEARKIWCFGPDGTGPNILTDITKG VQ----YLNE	673
Qy	474	VKDSIVQG	481
Db	674	IKDSVVAG	681

RESULT 7

EF2 CHICK

ID EF2 CHICK STANDARD; PRT; 857 AA.

AC Q90705;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Elongation factor 2 (EF-2).

GN EEF2.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI TaxID=9031;

CH	FILE
RN	[1]

RP SEQUENCE FROM N.A.

RC TISSUE=Intestine;

RA Kim C.W., Jung E.J., Ahn H.J., Kim J.C., Kang K.R., Eom M.-O.,
 RA Kim Y.W., Kang Y.-S.;
 RT "Molecular cloning of chicken elongation factor 2 (EF-2): sequence
 RT comparison with mammalian EF-2 and its expression in the early
 RT developmental stages of the embryos."
 RL Mol. Cells 3:27-33(1993).
 RN [2]
 RP SEQUENCE OF 1-21, AND PHOSPHORYLATION.
 RX MEDLINE=91207327; PubMed=1708237;
 RA Kim Y.W., Kim C.W., Kang K.R., Byun S.M., Kang Y.S.;
 RT "Elongation factor-2 in chick embryo is phosphorylated on tyrosine as
 RT well as serine and threonine."
 RL Biochem. Biophys. Res. Commun. 175:400-406(1991).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2
 CC (BY SIMILARITY).
 CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
 CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
 CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U46663; AAA87587.1; -.
 DR HSSP; P13551; 1FNM.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT INIT_MET 0 0
 FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 103 107 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 714 714 DIPHTHAMIDE (BY SIMILARITY).
 SQ SEQUENCE 857 AA; 95247 MW; 3680187581F518E6 CRC64;

Query Match 34.2%; Score 865; DB 1; Length 857;
 Best Local Similarity 37.5%; Pred. No. 9e-56;
 Matches 183; Conservative 94; Mismatches 189; Indels 22; Gaps 5;

```

QY      9 NLILSPLLGNVCFSSSQYSICFTLGSFAKIYADTF---GDI-----NYQEFARKLW 56
      |::: |:|| | | : || |::| | || : |:|
Db     201 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGDAQMNPTERAKKVEDMMKKLW 260

QY     57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGI 113
      || ||:| | ||:| | : |:| : ||:|::|: : : :::| |
Db    261 GDRYFDPATGKFSKATGPDGKKLPRTFCQLILDPIFKVFDAIMTFKKEEAAKLEKLDI 320

QY    114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
      | |: :||:| ::: : | :|:| | :| | | | :
Db    321 KLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEA 380

QY    174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
      : :|| | || : :|| | | :|: ||| | | : | |::| ||| :||
Db    381 AIGIKNCDPRGSLMMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKED 440

QY    234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
      : : | : : || : || || | : ||| :||| || | | | :|
Db    441 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMVRMK 497

QY    294 FNTTSVIKIAVEPVNPSELPMKLDGLRKVNKSYP SLTKVEESGEHVILGTGELYLDCVM 353
      |: : |:|:| | ||:|:|:|:|:|:| | | : :|:|:|:| | | | |:| :
Db    498 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIAGAGELHLEICL 557

QY    354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
      || : : : | | | :||:|:| | | | :| : :|:| | :| | | | | :
Db    558 KDLEEDHACIPIKKSDPVVSRET VSEESNMCLSKSPNKHNRLYMKARPPDGLAEDID 617

QY    414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
      | : : |:|:| | | | | | | | | | | | | : |
Db    618 KGEVSARQELKQRARYLAEKYEDVTEARKIWCFGPDGTGPNILTDITKGVQ----YLNE 673

QY    474 VKDSIVQG 481
      :||:| |
Db    674 IKDSVVAG 681
  
```

RESULT 8

EF2_MESAU

ID EF2_MESAU STANDARD; PRT; 857 AA.

AC P05086;

DT 13-AUG-1987 (Rel. 05, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Elongation factor 2 (EF-2).

GN EE2.

OS Mesocricetus auratus (Golden hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC Mesocricetus.

OX NCBI_TaxID=10036;

RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=86259716; PubMed=3014523;
 RA Kohno K., Uchida T., Ohkubo H., Nakanishi S., Nakanishi T.,
 RA Fukui T., Ohtsuka E., Ikehara M., Okada Y.;
 RT "Amino acid sequence of mammalian elongation factor 2 deduced from
 RT the cDNA sequence: homology with GTP-binding proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4978-4982(1986).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
 CC -!- PTM: HIS-714 IS MODIFIED IN DIPHTHAMIDE (2-[3-CARBOXYAMIDO-3-
 CC (TRIMETHYL-AMMONIO)PROPYL]HISTIDINE). DIPHTHAMIDE CAN BE ADP-
 CC RIBOSYLATED BY DIPHTERIA TOXIN AND BY PSEUDOMONAS EXOTOXIN A.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M13708; AAA50387.1; -.
 DR PIR; A25440; A25440.
 DR HSSP; P13551; 1FNM.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 103 107 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 714 714 DIPHTHAMIDE.
 SQ SEQUENCE 857 AA; 95192 MW; 42356544F43947AB CRC64;

Query Match 34.2%; Score 863; DB 1; Length 857;
 Best Local Similarity 37.1%; Pred. No. 1.3e-55;
 Matches 181; Conservative 95; Mismatches 190; Indels 22; Gaps 5;

Qy 9 NLILSPLLGNVCFSSSQYSICFTLGSAFKIYADTF-----GDI-----NYQEFAKRLW 56
 |::: |:|| | | : ||| |::| | : :: |:|

Db 201 NIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLW 260

QY 57 GDIYFNPKTRKFTKKAPTSSSQ---RSFVEFILEPLIKILAQVVGDVDTSLPRTLDELGI 113
 || ||:| ||:| | : : | : | : ||:|:|: : : : : || |

Db 261 GDRYFDPANGKFSKANSSPDGKKLPRTFCQLILDPIFKVFDAMNFRKEETAKLIEKLDI 320

QY 114 HLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS DL 173
 | | : : ||:| | : : : | | : || | | : | | | | :

Db 321 KLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLSPVTAQKYRCELLYEGPPDDEA 380

QY 174 GEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEED 233
 : ||| |||| : : || | | : | : |||| || : | | : || || || : | :

Db 381 AMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGVVSTGLKVRIMGPNYTPGKKEE 440

QY 234 SQICTVGRWLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLK 293
 : : | : : || : || || | : |||| : || | || | | | : |

Db 441 LYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITT---FEHAHNMRVMK 497

QY 294 FNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVM 353
 | : : | : || || || : || || : || : || : || : || : || : || : || : || :

Db 498 FSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCTIEESGEHIIAGAGELHLEICL 557

QY 354 HDLRKMYSEIDIKVADPVVTFCE TVVETSSSLKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
 || : : | || : |||| : || | | : | : || || : | | | | |||| :

Db 558 KDLEEDHACIPIKKS DPVVS YRET VSEES NVLCLSKSPNKHNR LYM KARPF PDGLAEDID 617

QY 414 NEVVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGS 473
 | | : ||:|: || || || || || || || || | : |

Db 618 KGEVSARQELKARARYLAEKYEW DVAEARKIWC FGP DGTGPNILT DITKGVQ---YLNE 673

QY 474 VKDSIVQG 481
 : ||| : | |

Db 674 IKDSVVAG 681

RESULT 9

EF2_CRYPV

ID EF2_CRYPV STANDARD; PRT; 832 AA.

AC Q23716;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Elongation factor 2 (EF-2).

OS Cryptosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

OC Cryptosporidiidae; Cryptosporidium.

OX NCBI_TaxID=5807;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AUCP-1;

RX MEDLINE=95356792; PubMed=7630379;

RA Jones D.E., Tu T.D., Mathur S., Sweeney R.W., Clark D.P.;

RT "Molecular cloning and characterization of a Cryptosporidium parvum

RT elongation factor-2 gene.";

RL Mol. Biochem. Parasitol. 71:143-147(1995).

CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION

CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE


```

CC      RIBOSOME (BY SIMILARITY).
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC      EF-G/EF-2 SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U21667; AAC46607.1; -.
DR      InterPro; IPR000795; EF_GTPbind.
DR      InterPro; IPR000640; EFG_C.
DR      InterPro; IPR005517; EFG_IV.
DR      InterPro; IPR004161; EFTU_D2.
DR      InterPro; IPR005225; Small_GTP.
DR      Pfam; PF00679; EFG_C; 1.
DR      Pfam; PF03764; EFG_IV; 1.
DR      Pfam; PF00009; GTP_EFTU; 1.
DR      Pfam; PF03144; GTP_EFTU_D2; 1.
DR      PRINTS; PR00315; ELONGATNFCT.
DR      TIGRFAMs; TIGR00231; small_GTP; 1.
DR      PROSITE; PS00301; EFACTOR_GTP; 1.
KW      Elongation factor; Protein biosynthesis; GTP-binding; Phosphorylation.
FT      NP_BIND      26      33      GTP (BY SIMILARITY).
FT      NP_BIND      98     102      GTP (BY SIMILARITY).
FT      NP_BIND     152     155      GTP (BY SIMILARITY).
FT      MOD_RES      57      57      PHOSPHORYLATION (BY SIMILARITY).
FT      MOD_RES      59      59      PHOSPHORYLATION (BY SIMILARITY).
FT      MOD_RES     689     689      DIPHTHAMIDE (BY SIMILARITY).
SQ      SEQUENCE    832 AA;  92761 MW;  577D2DE23D77E3FA CRC64;

```

```

Query Match      33.8%;  Score 854;  DB 1;  Length 832;
Best Local Similarity 37.3%;  Pred. No. 5.5e-55;
Matches 174;  Conservative 99;  Mismatches 175;  Indels 18;  Gaps 4;

```

```

QY      9  NLILS-----PLLGNVCFSSSQYSICFTLGSF AKIYADTFGDINYQEFAKRL 55
      |::| | | | | : |:: ||: || | | | : : ||
Db      180 NVIISTYSDELMGDVQVFPEKGTVSFGSLHGWAFTEKFARIYAKKFG-VEKSKMMQRL 238

QY      56  WGDIYFNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHL 115
      ||| :|||:|:| | | | :|:| :||:|: : : : : : | | : |
Db      239 WGNFNFNPETKKFT-KTQEPGSKRAFCQFIMEPICQLFSSIMNGDKAKYEKMLVNLGVEL 297

QY      116 TKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTG GVDSDLGE 175
      :: | :||: | : : : :| | | :|| | :|: | | | : :
Db      298 KGDDKALVDKPLLKKVQLWLSAGDTLLEMIVTHLPSAAAQKYRVENLYEGPQDDETAK 357

QY      176 AMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQ 235
      : :||| | | :|| | | :|: ||| ||: | | | :| | :||
Db      358 GIRNCDPDAPLCMFVSKMVPTSDKGRFYAFGRVFGT VATGQKVRIQGPRYVPGGKEDLN 417

QY      236 ICTVGRWLISVARYHIEVNRVPAGNWWLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFN 295
      | : | : : || : : |||| | : |:| :|: | | | :| | :|:

```

Db 418 IKNIQRTVLMMGRYVEQIPDVPAGNTVGLVGIDQYLLKSGTITTT---SETAHNIASMKYS 474

Qy 296 TTSVIKIAVEPVNPSELPMKLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHD 355
: |::|| | : |||:::|::|| | : ||:|::| | |||::: : |

Db 475 VSPVVRVAVRPKDNKELPKLVEGLKKLSKSDPLVCSKEETGEHIIAGCGELHVEICLQD 534

Qy 356 LRKMYSEIDIKVADPVVTFCETVVETSSLKCAETPNKKNKITMIAEPLEKGLAEDIENE 415
|:: |::|:| :||:|:: ||| | : | ::||| |:: | ||| | | :|||

Db 535 LQQEYAQIEIVASDPIVSYRETVVNLSNQTCLSKSPNKHNRLYMTAEPLPDGLTDDIEEG 594

Qy 416 VVQITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDT 461
| : | : || : | || || |||: ||||:| | |

Db 595 KVSRRDDPKERSNLLHDKYGFDKNAAMKIWCFGPETTGPNNIMVDVT 640

RESULT 10

EF2_DROME

ID EF2_DROME STANDARD; PRT; 843 AA.

AC P13060; Q9I7H2; Q9V9R0;

DT 01-JAN-1990 (Rel. 13, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Elongation factor 2 (EF-2).

GN EF2B OR CG2238.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RX MEDLINE=90016792; PubMed=2508059;

RA Grinblat Y., Brown N.H., Kafatos F.C.;

RT "Isolation and characterization of the Drosophila translational

RT elongation factor 2 gene.";

RL Nucleic Acids Res. 17:7303-7314(1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";

RL Science 287:2185-2195(2000).

CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION OF
 CC THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=Long;

CC IsoId=Pl3060-1; Sequence=Displayed;

CC Name=Short;

CC IsoId=Pl3060-2; Sequence=VSP_001361, VSP_001362;

CC Note=No experimental confirmation available;

CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.

CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X15805; CAA33804.1; -.

DR EMBL; AE003781; AAF57226.1; ALT_SEQ.

DR EMBL; AE003781; AAG22125.1; ALT_SEQ.

DR PIR; S05988; S05988.

DR FlyBase; FBgn0000559; Ef2b.

DR InterPro; IPR000795; EF_GTPbind.

DR InterPro; IPR000640; EFG_C.

DR InterPro; IPR005517; EFG_IV.

DR InterPro; IPR004161; EFTU_D2.

DR InterPro; IPR005225; Small_GTP.

DR Pfam; PF00679; EFG_C; 1.

DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation;
 KW Alternative splicing.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT NP_BIND 25 32 GTP (BY SIMILARITY).
 FT NP_BIND 107 111 GTP (BY SIMILARITY).
 FT NP_BIND 161 164 GTP (BY SIMILARITY).
 FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 700 700 DIPHTHAMIDE (BY SIMILARITY).
 FT VARSPLIC 700 710 HRGGGQIIPTT -> TSKCSWLSGKA (in isoform
 FT Short).
 FT /FTid=VSP_001361.
 FT VARSPLIC 711 843 Missing (in isoform Short).
 FT /FTid=VSP_001362.
 FT CONFLICT 606 606 D -> E (IN REF. 1).
 SQ SEQUENCE 843 AA; 94327 MW; C4BEE437F23E5593 CRC64;

Query Match 33.4%; Score 844.5; DB 1; Length 843;
 Best Local Similarity 37.2%; Pred. No. 2.8e-54;
 Matches 175; Conservative 103; Mismatches 183; Indels 9; Gaps 4;

QY 12 LSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFAKRLWGDIFYNPKTRKFTKK 71
 : | | : | | : ||| |::|:: | | : : |||| : || ||:| : |
 Db 207 VDPKSGSVGFSGGLHGWAFTLKQFSEMYSEKF-KIDVVKLMNRLWGENFFNAKTKKW-QK 264
 QY 72 APTSSSQRSFVEFILEPLYKILAQVVGVDVDTSLPRTLDELGIHLTKHEELKLNIRPLRLV 131
 : ::||| : ||:|::| : : : |::|: | | : : ||: |
 Db 265 QKEADNKRSCMYILDPIYKVFDAIMNYKKEEIGTLLEKIGVTLKHEDKDKDGKALLKTV 324
 QY 132 CKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHTT 191
 : : : | |::| | | : | | | : | : ||||| : :
 Db 325 MRTWLPAGEALLQMIAIHLPSPVVAQKYRMEMLYEGPHDDEAAIAVKSCDPDGPLMMYIS 384
 QY 192 KMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRLWISVARYHI 251
 || | | :|: ||| :| : || :::| ||| :|| : | : : ||
 Db 385 KMVPTSDKGRFYAFGRVFAGKVATGQKCRIMGPNYTPGKKEDLYEKAIQRTILMMGRYVE 444
 QY 252 EVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPSE 311
 : ||:| : ||| :|| || :| : :||: : |::||| ||::
 Db 445 AIEDVPSGNICGLVGVDQFLVKTGTITT---FKDAHNMKVMKFSVSPVVRVAVEPKNEAD 501
 QY 312 LPKMLDGLRKVNKSYPSTTKVEESGEHVILGTGELYLDVCMHDLRKMYSIDIKVADPV 371
 |||::|::|:: || | : :|||::| | |||:| : || : : | :| :|||
 Db 502 LPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPLKSDPV 561
 QY 372 VTFCETVVETSSSLKCFEAETPNKKNKITMIAEPLKGLAEDIENEVVQITWNRKKLGFEFFQ 431
 |:: ||| | | :::|| |:: | | | | |||:| | : :
 Db 562 VSYRETVSEESDQMCLSKSPNKHNRLLMKALPMPDGLPEDIDNGDVSADKDEFKARARYLS 621
 QY 432 TKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSKVKSIVQG 481
 |||:| : || || ||| ||| :| | : | :||:| |

Db

622 EKYDYDVTEARKIWCFPGDGTGPNFILDCTKSVQ-----YLNEIKDSVVAG 667

RESULT 11

EF2_CANAL

ID EF2_CANAL STANDARD; PRT; 842 AA.
AC O13430;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor 2 (EF-2).
GN EFT2.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RA Capa L., Mendoza A., Serramia M.J., Garcia-Bustos J.F.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y09664; CAA70857.2; -.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACTOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 26 33 GTP (BY SIMILARITY).
FT NP_BIND 104 108 GTP (BY SIMILARITY).
FT NP_BIND 158 161 GTP (BY SIMILARITY).
FT MOD_RES 699 699 DIPHTHAMIDE (BY SIMILARITY).
SQ SEQUENCE 842 AA; 93354 MW; 5C8C740109A23189 CRC64;

Query Match

33.2%; Score 839; DB 1; Length 842;

Best Local Similarity 35.6%; Pred. No. 7.1e-54;
Matches 174; Conservative 116; Mismatches 171; Indels 28; Gaps 6;

```

Qy      9 NLILS-----PLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFKRL 55
      |:|:| | | | |:| : ||: || |: || :: :: :||
Db     186 NVIISTYCDPVLGDVQVYPQKGTVAFAASGLHGWAFTVRQFANKYSKKFG-VDKEKMMERL 244

Qy      56 WGDIYFNPKTRKFTKK---APTSSSQRSFVEFILEPLYKILAQVVGDDVDSLPRTLDELG 112
      ||| |||||:|:| | | :|:| |||:|::: | :: :| |:|
Db     245 WGDSYFNPKTKKWTNKDKDADGKPLERAFNMFILDPIFRLFAAIMNFKKDEIPVLLLEKLE 304

Qy     113 IHLTKEELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVDS 172
      | | :| | : ||:| :|| :| | | | | : | | | |
Db     305 IQLKGDEKDLEGKALLKVMRKFLPAADALLEMIVLHLPSPVTAQAYRAETLYEGPSDDP 364

Qy     173 LGEAMSDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYLTLEDEE 232
      |: :||: || : :|| | | :|: ||| :|: :|| |:| | | : :|
Db     365 FCTAIRNCDPNADLMLYVSKMVPSTDKGRFYAFGRVFAGTVKSGQKVRIQGPNYQVGKKE 424

Qy     233 DSQICTVGRWLWISVARYHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPL 292
      | : :: | : : | ::| ||| : : |:| :|: ||| || | : :
Db     425 DLFLKSIQRTVLMMGRSVEQIDDCPAGNIIGLVGIDQFLLKSGTITT---NEAAHNMKVM 481

Qy     293 KFNTTSVIKIAVEPVNPSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCV 352
      ||: : |:|:| | :|:|:|:|:|:| | | : | | | :| | | :| :|
Db     482 KFSVSPVVQVAVEVKNANDLPKLIVEGLKRLSKSDPCVLTYSSESGEHIVAGTGELHLEIC 541

Qy     353 MHDLRKMYSEIDIKVADPVVTFCE TVVETSSLKCF AETPNKKNKITMIAEPLEKGLAEDI 412
      : || : : : ::| |||: ||| ||: :|:| | |:| : |:|: : : ||
Db     542 LQDLENDHAGVPLRISPPVVS YRETVEGESSMVALSKSPNKHNRIVKAQPIDEEVSLDI 601

Qy     413 ENEVVQITWNRKKLGFEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTL PSEVDKAL-- 470
      || |: : | :|:|:| || || ||| |||:| | | | | :
Db     602 ENGVINPRDDFKARARILADKHGWDVVDARKIWCFGPDGNGPNLVVDQT-----KAVQY 655

Qy     471 LGSVKDSIV 479
      | :||:|
Db     656 LNEIKDSVV 664

```

RESULT 12

EF2_CAEEL

ID EF2_CAEEL STANDARD; PRT; 851 AA.
AC P29691; O17837;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor 2 (EF-2).
GN EFT-2 OR F25H5.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92029622; PubMed=1930695;
RA Ofulue E.N., Candido E.P.M.;

RT "Molecular cloning and characterization of the *Caenorhabditis elegans*
RT elongation factor 2 gene (*eft-2*).";
RL DNA Cell Biol. 10:603-611(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Steward C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M86959; AAD03339.1; -.
DR EMBL; Z81068; CAB02985.1; -.
DR PIR; A40411; A40411.
DR PIR; T21362; T21362.
DR WormPep; F25H5.4; CE15900.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT NP_BIND 25 32 GTP (BY SIMILARITY).
FT NP_BIND 115 119 GTP (BY SIMILARITY).
FT NP_BIND 169 172 GTP (BY SIMILARITY).
FT MOD_RES 56 56 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 58 58 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 708 708 DIPHTHAMIDE (BY SIMILARITY).
FT CONFLICT 131 131 R -> G (IN REF. 1).
FT CONFLICT 616 616 N -> S (IN REF. 1).
FT CONFLICT 627 629 ILA -> YPG (IN REF. 1).
FT CONFLICT 635 635 D -> A (IN REF. 1).
SQ SEQUENCE 851 AA; 94665 MW; 2B2460CEB6399176 CRC64;

Query Match 33.1%; Score 837.5; DB 1; Length 851;
Best Local Similarity 36.2%; Pred. No. 9.3e-54;
Matches 175; Conservative 103; Mismatches 193; Indels 13; Gaps 5;

Qy	2	AYYSTDEN----LILSPLLGNVCFSSSQYSICFTLGSFAKIYADTFGDINYQEFAKRLWG	57
Db	201	ATYGDDDGPMGPIMVDPSIGNVGFGSGLHGWAFTLQKQFAEMYAGKFG-VQVDKLMKNLWG	259
Qy	58	DIYFNPKTRKFTKKAPTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTK	117
Db	260	DRFFDLKTKKWS-STQTDESKRGFCQFVLDPIFMVFDAMVNIKKDKTAALVEKLGLIKLAN	318
Qy	118	EELKLNIRPLLRLVCKKFFGEFTGFVDMCVQHIPSPKVGAKPKIEHTYTGGVDSDLGEAM	177
Db	319	DEKDLEGKPLMKVEMRKWLPAAGDTMLQMIAFHLPSPVTAQKYRMEMLYEGPHDDEAAVAI	378
Qy	178	SDCDPDGPLMCHTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTTLEDEEDSQIC	237
Db	379	KTCDPNGPLMMYISKMVPTSDKGRFYAFGRVFSGKVATGMKARIQGPNYVPGKKEDLYEK	438
Qy	238	TVGRLWISVARYHIEVNRVPAGNWWLIEGVDQPIVKTTATITEPRGNEEAQIFRPLKFNTT	297
Db	439	TIQRTILMMGRFIEPIEDIPSGNIAGLVGVDQYLVKGGTITT---YKDAHNMVRVMKFSVS	495
Qy	298	SVIKIAVEPVNPSELPMKMLDGLRKVNKSYPSTLTKVEESGEHVILGTGELYLDCVMHDLR	357
Db	496	PVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIFEESGEHIIAGAGELHLEICLKDLE	555
Qy	358	KMYSEIDIKVADPVVTFCETVVETSSSLKCFETPNKKNKITMIAEPLEKGLAEDIENEV	417
Db	556	EDHACIPLKKSDFVVSYRETQVSESNQICLSKSPNKHNRHLCTAQPMPDGLADDIEGGTV	615
Qy	418	QITWNRKKLGEFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKALLGSVKDS	477
Db	616	NARDEFKARAKILAEKYEYDVTEARKIWCFGPDGTGPNLLMDVTKGVQ----YLNEIKDS	671
Qy	478	IVQG 481	
Db	672	VVAG 675	

RN [2]
 RP SEQUENCE OF 30-790 FROM N.A.
 RC STRAIN=HK-9;
 RX MEDLINE=94355065; PubMed=8074887;
 RA Shirakura T., Hashimoto T., Nakamura Y., Kamaishi T., Cao Y.,
 RA Adachi J., Hasegawa M., Yamamoto A., Goto N.;
 RT "Phylogenetic place of a mitochondria-lacking protozoan, *Entamoeba*
 RT *histolytica*, inferred from amino acid sequences of elongation factor
 RT 2'.";
 RL Jpn. J. Genet. 69:119-135(1994).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L02417; AAA29097.1; -.
 DR EMBL; D21259; BAA04800.1; -.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFATOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation.
 FT NP_BIND 26 33 GTP (BY SIMILARITY).
 FT NP_BIND 102 106 GTP (BY SIMILARITY).
 FT NP_BIND 156 159 GTP (BY SIMILARITY).
 FT MOD_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 59 59 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 697 697 DIPHTHAMIDE (BY SIMILARITY).
 FT CONFLICT 471 471 V -> DT (IN REF. 2).
 FT CONFLICT 541 541 E -> D (IN REF. 2).
 FT CONFLICT 592 592 P -> A (IN REF. 2).
 SQ SEQUENCE 840 AA; 93324 MW; 2E705327E3B52210 CRC64;

Query Match 32.6%; Score 823.5; DB 1; Length 840;
 Best Local Similarity 36.3%; Pred. No. 9.8e-53;
 Matches 173; Conservative 104; Mismatches 182; Indels 17; Gaps 6;

QY 12 LSPLLGNVCFSSSQYSICFTLGSEFAKIYADTFGDINYQEFARKLWGDYFNPKTRKFTKK 71

```

Db      200  :|| | | | : ||| |||::: || |: : :||| |:: |:: |
VSPGEGTVAFGSLHGWAFMLEKFAKMWSAKFG-IDRKRMLEKLWGDNYWDAAKAKWKKN 258

Qy      72  APTSSS---QRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKLNIRPLL 128
|| ||:| :|: | : : | | | : : : : ||
Db      259  GKGDHGEVLQRGFVQFCFDPITKLFNAIMEGRKADYEKMLTNLQIKLSADDKEKEGKELL 318

Qy      129  RLVCKKFFGEFTGFVDMCVQHIPS PKVGAKPKIEHTYTGGVSDLGEMS DCDPDGPLMC 188
: | | : :|| | |::| | | : : ||| : | : :||::|| |||
Db      319  KTVMKLWLPAGVTLLMIVLHLPSPVVAQKYRTSNLYTGPMDEAAKAMANCDEKGPLMM 378

Qy      189  HTTKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDS QICTVGRWLWISVAR 248
: || | :| :|: ||| ||| | : : | || :|| | : | : : |
Db      379  YVSKMIPTNDKGRFYAFGRVFSGTIRTGGKARICGPNYVPGKDDCVIKNIQRTMLMMGR 438

Qy      249  YHIEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVN 308
| : : | | : : ||| :|: ||| : | | : :||: : |:::|| |
Db      439  YTDPIDECPCGNVIGLVGDVQYLLKSGTITD-----SVAHIKDMKFSVSPVVRVAVETKN 494

Qy      309  PSELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMV-SEIDIKV 367
||:||||::|:|:::| | |||||:: | |||:| : :|:: | | : : |
Db      495  PSDLPKLVEGMKRLSRSDPLCLCYTEESGEHIVAGAGELHLEVCLKELQEDYCSGVPLIV 554

Qy      368  ADPVVTFCETVVETSSSLKCFEAETPNKKNKITMIAEPELEKGLAEDIENEVVQITWNRKKLG 427
: |||:| ||: | | ::| ::| :|: | | | :||| ||| : : | :
Db      555  TEPVVSFRETITEPSRIQCLSKSANNQNRLFMRAFPFPEGLAEDIEAGEIKPDTDFKERA 614

Qy      428  EFFQTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIVQG 481
: | || ||: || || ||| |||: || | | : | ||||| |
Db      615  KFLSEKYGWDVDDEARKIWCFGPDNCGPNLFVD-----VTKGIQYLNEVKDSIVNG 664

```

RESULT 14

EF2_BETVU

```

ID  EF2_BETVU      STANDARD;      PRT;      843 AA.
AC  023755;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Elongation factor 2 (EF-2).
OS  Beta vulgaris (Sugar beet).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX  NCBI_TaxID=161934;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Root;
RA  Vogel R., Rausch T.;
RL  Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC  -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC      OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC      RIBOSOME.
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- PTM: PHOSPHORYLATION BY EF-2 KINASE COMPLETELY INACTIVATES EF-2
CC      (BY SIMILARITY).
CC  -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.

```


Db 484 SVSPVVRVAVQCKVASDLPKLV EGLKRLAKSDPMVVC SIEESGEHI IAGAGELHLEICLK 543
 QY 355 DLRKMY-SEIDIKVADPVVTF CETVVETSS LKCF AETPNKKNKITMIAE PLEKGLAEDIE 413
 ||: : :| :|||:| |||:| :||| :| :| :|:|||| |:|
 Db 544 DLQDDFMGGAEIIKSDPVVS FRETVLDRSVRTVMSKSPNKHNR LYM EARPMEEG LAE AID 603
 QY 414 NEVVQITWNRKKLGEFFQTKYDWDLLA ARSIWAFGPDATGPNILVDDTLPSEVDKAL--L 471
 : :| : :| || |: || |||: |||:| | :| :|
 Db 604 EGRIGPRDDPKNRSKILAE EYGWDKDLAKKIWC FGPETTGP NMVVD-----MCKGVQYL 657
 QY 472 GSVKDSIVQG 481
 :|||:| |
 Db 658 NEIKDSVVAG 667

RESULT 15

EF2_YEAST

ID EF2_YEAST STANDARD; PRT; 842 AA.
 AC P32324;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Elongation factor 2 (EF-2).
 GN (EFT1 OR YOR133W OR O3317 OR YOR3317W) AND (EFT2 OR YDR385W).
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. (EFT1 AND EFT2).
 RX MEDLINE=92112760; PubMed=1730643;
 RA Perentesis J.P., Phan L.D., Laporte D.C., Livingston D.M.,
 RA Bodley J.W.;
 RT "Saccharomyces cerevisiae elongation factor 2. Genetic cloning,
 RT characterization of expression, and G-domain modeling."
 RL J. Biol. Chem. 267:1190-1197(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (EFT1).
 RX MEDLINE=97344368; PubMed=9200815;
 RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
 RA Schwager C., Paces V., Sander C., Ansorge W.;
 RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV."
 RL Yeast 13:655-672(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (EFT1).
 RC STRAIN=S288c / FY1679;
 RX MEDLINE=97060020; PubMed=8904341;
 RA Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vlcek C.,
 RA Stegemann J., Zimmermann J., Erfle H., Paces V., Ansorge W.;
 RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV
 RT from *Saccharomyces cerevisiae* reveals 30 open reading frames."
 RL Yeast 12:281-288(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (EFT2).
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
 RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,

RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE OF 411-422 AND 505-513.
 RC STRAIN=S288c;
 RX MEDLINE=95203288; PubMed=7895733;
 RA Garrels J.I., Fitcher B., Kobayashi R., Latter G.I., Schwender B.,
 RA Volpe T., Warner J.R., McLaughlin C.S.;
 RT "Protein identifications for a *Saccharomyces cerevisiae* protein
 RT database.";
 RL Electrophoresis 15:1466-1486(1994).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M59369; AAA21646.1; -.
 DR EMBL; M59370; AAA51398.1; -.
 DR EMBL; X90518; CAA62116.1; -.
 DR EMBL; X94335; CAA64052.1; -.
 DR EMBL; Z75041; CAA99332.1; -.
 DR EMBL; U32274; AAB64827.1; -.
 DR EMBL; U28373; AAB64821.1; -.
 DR PIR; A41778; A41778.
 DR PDB; 1NOU; 11-FEB-03.
 DR SWISS-2DPAGE; P32324; YEAST.
 DR SGD; S0005659; EFT1.
 DR SGD; S0002793; EFT2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR000640; EFG_C.
 DR InterPro; IPR005517; EFG_IV.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EFG_C; 1.
 DR Pfam; PF03764; EFG_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACTOR_GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis; Phosphorylation;
 KW 3D-structure.
 FT NP_BIND 26 33 GTP (BY SIMILARITY).
 FT NP_BIND 104 108 GTP (BY SIMILARITY).
 FT NP_BIND 158 161 GTP (BY SIMILARITY).

FT MOD_RES 699 699 DIPHTHAMIDE (BY SIMILARITY).
SQ SEQUENCE 842 AA; 93289 MW; FD2F8073CB9B66AA CRC64;

Query Match 32.0%; Score 807.5; DB 1; Length 842;
Best Local Similarity 35.7%; Pred. No. 1.5e-51;
Matches 168; Conservative 106; Mismatches 182; Indels 15; Gaps 5;

```
Qy      14 PLLGNVCFSSSQYSICFTLGSAFKIYADTFGDINYQEFARLWGDIFYNPKTRKFTKK-- 71
      | | | | | : ||: || || || :: : |||| :|||:|:| |
Db      204 PARGTVAFGSLHGWAFTIRQFATRYAKKFG-VDKAKMDRLWGDSFFNPKTKKWTNKDT 262

Qy      72 -APTSSSQRSFVEFILEPLYKILAQVVGDVDTSLPRTLDELGIHLTKEELKLNIRPLLRL 130
      | :|:| |||:|::: : : | |::| | | :| | : ||::
Db      263 DAEGKPLERAFNMFILDPIFRLFTAIMNFKKDEIPVLEKLEIVLKGDEKDLEGKALLKV 322

Qy     131 VCKKFFGEFTGFVDMCVQHIPSQVGAQPKIEHTYTGGVSDSLGEAMSDCDPDGPLMCHT 190
      | :|| :|:| |||:| :| | | | | :| :||| || :
Db     323 VMRKFLPAADALLEMIVLHLPSPVTAQAYRAEQLYEGPADDANCIAIKNCDPKADLMLYV 382

Qy     191 TKMFSTHDGVQFHPFGRVLSGTIHAGQPVKVLGENYTLDEEDSQICTVGRWLWISVARYH 250
      :|| | | :|: ||| :||: || |:: | | :| | : | : : | :
Db     383 SKMVPTSDKGRFYAFGRVFAGTVKSGQKVRIQGPNYVPGKKDDLFIKAIQRVVLMGRFV 442

Qy     251 IEVNRVPAGNWVLIEGVDQPIVKTATITEPRGNEEAQIFRPLKFNTTSVIKIAVEPVNPS 310
      :: ||| : : |::| ::| |::| :| | : :||: : |::||| | :
Db     443 EPIDDCPAGNIIGLVGIDQFLLKTGTLTT---SETAHNMKVMKFSVSPVVQVAVEVKNAN 499

Qy     311 ELPKMLDGLRKVNKSYPSLTTKVEESGEHVILGTGELYLDCVMHDLRKMYSIDIKVADP 370
      :|||::|:|:|:| | : | : ||||:|: ||||:|: : || : : :|:| |
Db     500 DLPKLVEGLKRLSKSDPCVLTYMSESGEHIVAGTGELHLEICLDLEHDHAGVPLKISPP 559

Qy     371 VVTFCETVVETSSLKCAETPNKKNKITMIAEPEKGLAEDIENEVVQITWNRKKLGEFF 430
      || : ||| || :::||| |::| : ||::: : : ||| : : : |
Db     560 VVAYRETVESESSQTALSKSPNKHNRILYKAEPIDEEVSLAIENGINPRDDFKARARIM 619

Qy     431 QTKYDWDLLAARSIWAFGPDATGPNILVDDTLPSEVDKAL--LGSVKDSIV 479
      | ||: || || ||| ||:::| | ||: | :||:|
Db     620 ADDYGWDVTDARKIWCFGPDGNGPNLVIDQT-----KAVQYLHEIKDSVV 664
```

Search completed: January 30, 2004, 11:25:06
Job time : 19.716 secs